Introduction to Geometric Measure Theory

Leon Simon [1]

[1] The research described here was partially supported by NSF grants DMS-9504456 & DMS–9207704 at Stanford University

Contents

Preface to the Tsinghua Lectures 2014

The present text is a revision and updating of the author's 1983 "Lectures on Geometric Measure Theory," and is meant to provide an introduction to the subject at beginning/intermediate graduate level. The present draft is still in rather rough form, with a generous scattering of (hopefully not serious, mainly expository) errors. During the Tsinghua lectures (February–April 2014) the notes will be further revised, with the ultimate aim of providing a useful and accessible introduction to the subject at the appropriate level.

The author would greatly appreciate feedback about errors and other deficiencies.

Leon Simon
Beijing, China, February 2014
lsimon@stanford.edu

Last updated: Wed Dec 3 at 3:05pm

Notation

$\overline{A}$ = closure of A, assuming A is a subset of some topological space X

$B \setminus A = \{x \in B : x \notin A\}$

χ_A = indicator function of A (= 1 at points of A and = 0 and points not in A)

I_A = identity map $A \to A$

$\mathcal{L}^n$ = Lebesgue outer measure in $\mathbb{R}^n$

$B_\rho(y)$ = closed ball with center y radius ρ (more specifically denoted $B^n_\rho(y)$ if we wish to emphasize that we are working in $\mathbb{R}^n$). Thus $B_\rho(y) = \{x \in \mathbb{R}^n : |x-y| \le \rho\}$, or more generally, in any metric space X, $B_\rho(y) = \{x \in X : d(x,y) \le \rho\}$.

$\breve{B}_\rho(y)$ = open ball $= \{x \in \mathbb{R}^n : |x-y| < \rho\}$;

$\omega_k = \dfrac{\pi^{k/2}}{\int_0^\infty t^{k/2}e^{-t}\,dt}$ for $k \ge 0$ (so $\omega_k = \mathcal{L}^k(\{x \in \mathbb{R}^k : |x| \le 1\})$ if $k \in \{1,2,\ldots\}$).

$\eta_{y,\lambda} : \mathbb{R}^n \to \mathbb{R}^n$ (for $\lambda > 0$, $y \in \mathbb{R}^n$) is defined by $\eta_{y,\lambda}(x) = \lambda^{-1}(x-y)$; thus $\eta_{y,1}$ is translation $x \mapsto x - y$, and $\eta_{0,\lambda}$ is homothety $x \mapsto \lambda^{-1}x$

$C^k(U,V)$ (U, V open subsets of Euclidean spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively) denotes the space of C^k maps from U into V

For $f \in C^1(U,V)$, Df is the derivative matrix with entries $D_i f_j$ and the i-th row and j-th column, and $|Df|^2 = \sum_{i=1}^n \sum_{j=1}^m (D_i f_j)^2$.

For $f \in C^1(U,\mathbb{R})$ with U open in $\mathbb{R}^n$, we write $\nabla f = (D_1 f, \ldots, D_n f)\ (= (Df)^T)$.

$C^k_c(U,V) = \{\varphi \in C^k(U,V) : \varphi \text{ has compact support}\}$

For an abstract set X, 2^X denotes the collection of all subsets of X

$\emptyset$ = the empty set.

For any set A in a metric space X with metric d, diam A denotes the diameter of the set A, i.e. $\sup_{x,y\in A} d(x,y)$, interpreted to be zero if A is empty and ∞ if A is not bounded.

For $\Omega \subset \mathbb{R}^n$ open, $W^{1,2}(\Omega)$ will denote the Sobolev space of functions $f : \Omega \to \mathbb{R}$ such that $f, \nabla f \in L^2(\Omega)$.

If $W \subset U$, U open in $\mathbb{R}^n$, $W \subset\subset U$ means $\overline{W}$ is a compact subset of U.

δ_{ij} = Kronecker delta (= 1 if $i = j$, 0 if $i \ne j$).

Chapter 1

Preliminary Measure Theory

In this chapter we briefly review the basic theory of outer measure, which is based on Caratheodory's definition of measurability. Hausdorff (outer) measure is discussed, including the main results concerning n-dimensional densities and the way in which they relate more general measures to Hausdorff measures. The final two sections of the chapter give the basic theory of Radon (outer) measures including the Riesz representation theorem and the standard differentiation theory for Radon measures.

For the first section of the chapter X will denote an abstract space, and later we impose further restrictions on X as appropriate. For example in the second and third sections X is a metric space and in the last section of the chapter we shall assume that X is a locally compact, separable metric space.

1 Basic Notions

Recall that an outer measure (sometimes simply called a *measure* if no confusion is likely to arise) on X is a monotone subadditive function $\mu : 2^X \to [0, \infty]$ with $\mu(\emptyset) = 0$. Thus $\mu(\emptyset) = 0$, and

1.1 $$\mu(A) \le \sum_{j=1}^{\infty} \mu(A_j) \quad \text{whenever} \quad A \subset \cup_{j=1}^{\infty} A_j$$

with A, A_1, A_2, ... any countable collection of subsets of X. Of course this in particular implies $\mu(A) \le \mu(B)$ whenever $A \subset B \subset X$.

We adopt Caratheodory's notion of measurability:

1.2 Definition: A subset $A \subset X$ is said to be μ-measurable if

$$\mu(S) = \mu(S \setminus A) + \mu(S \cap A)$$

for each subset $S \subset X$. (Thus, roughly speaking, A is μ-measurable if it "cuts every other set S additively with respect to μ.")

Since $X \setminus (X \setminus A) = A$ we see that μ-measurability of A is equivalent to μ-measurability of $X \setminus A$ for any set $A \subset X$.

1.3 Remark: Then the set A is μ-measurable if and only if

$$\mu(S) \ge \mu(S \setminus A) + \mu(S \cap A)$$

for each subset $S \subset X$ with $\mu(S) < \infty$, because this is trivially true when $\mu(S) = \infty$, and the reverse inequality also holds in both cases $\mu(S) < \infty$ and $\mu(S) = \infty$ by the subadditivity 1.1 of μ.

Notice that the empty set $\emptyset$ is μ-measurable, as is any set of μ-measure zero since in this case the term $\mu(S \cap A)$ on the right side of the inequality in Remark 1.3 is zero.

A key lemma, due to Caratheodory, asserts that such μ-measurable sets form a σ-algebra, where the terminology is as follows:

1.4 Definition: A collection $\mathcal{S}$ of subsets of X is a *σ-algebra* if

(1) $\emptyset, X \in \mathcal{S}$

(2) $A \in \mathcal{S} \Rightarrow X \setminus A \in S$

(3) $A_1, A_2, \ldots \in \mathcal{S} \Rightarrow \cup_{j=1}^{\infty} A_j \in \mathcal{S}$.

1.5 Remarks: (1) Observe that then $\cap_{j=1}^{\infty} A_j = X \setminus \big(X \setminus (\cup_{j=1}^{\infty} A_j)\big) \in \mathcal{S}$ whenever $A_1, A_2, \ldots \in \mathcal{S}$, by (2), (3).

(2) In the context of a fixed space X, it is easily checked that the intersection of any non-empty family of σ-algebras is again a σ-algebra, so there is always a smallest σ-algebra which contains a given collection of subsets of X—namely just take the intersection of all σ-algebras which contain the given collection of sets (this collection is non-empty because the collection of all subsets of X is a σ-algebra). In particular if X is a topological space then there is a smallest σ-algebra containing all the open sets (same as the smallest σ-algebra containing all the closed sets since σ-algebras are closed under complementation), referred to as *"the Borel sets in X."*

As mentioned above, we have the following lemma:

1.6 Lemma. *The collection $\mathcal{M}$ of all μ-measurable subsets is a σ-algebra which includes all subsets of X of μ-measure zero.*

1.7 Remark: In the course of the proof we shall establish the important additional fact that for μ-measurable sets A_j, $j = 1, 2, \ldots$,

$$A_1, A_2, \ldots \text{ pairwise disjoint } \Rightarrow \mu(S \cap (\cup_{j=1}^{\infty} A_j)) = \textstyle\sum_{j=1}^{\infty} \mu(S \cap A_j)$$

for each subset $S \subset X$.

Proof of Lemma 1.6 and Remark 1.7: We already noted above that Properties 1.4(1) and 1.4(2) are trivially checked direct from the definition of measurability.

Checking 1.4(3) involves several steps:

Step 1: $A_1, A_2 \in \mathcal{M} \Rightarrow A_2 \setminus A_1 \in \mathcal{M}$ (which is just 1.4(2) in case $A_2 = X$). To check this we first use Definition 1.2 with $A = A_1$ and with $S \cap A_2$ in place of S and $\mu(S) < \infty$ to give

$$\mu(S \cap (A_2 \setminus A_1)) = \mu(S \cap A_2) - \mu(S \cap A_2 \cap A_1)$$

and then use Definition 1.2 with $A = A_2$ on the right side to give

$$\begin{aligned} \mu(S \cap (A_2 \setminus A_1)) &= \mu(S) - \mu(S \setminus A_2) - \mu(S \cap A_2 \cap A_1) \\ &\le \mu(S) - \mu((S \setminus A_2) \cup (S \cap A_2 \cap A_1)) \\ &= \mu(S) - \mu((S \setminus A_2) \cup (S \cap A_1)) \\ &= \mu(S) - \mu(S \setminus (A_2 \setminus A_1)) \end{aligned}$$

so $A_2 \setminus A_1$ is μ-measurable by Remark 1.3.

step 2: $A_1, A_2 \in \mathcal{M} \Rightarrow A_1 \cup A_2 \in \mathcal{M}$. To check this we simply observe that $(X \setminus A_1) \cap (X \setminus A_2) \in \mathcal{M}$ by using Step 1 with $X \setminus A_2 (\in \mathcal{M})$ in place of A_2. Then $A_1 \cup A_2 = X \setminus ((X \setminus A_1) \cap (X \setminus A_2)) \in \mathcal{M}$, so Step 2 is proved. Notice that if A_1, A_2 are disjoint then we also have the additional additivity conclusion that $\mu(S \cap (A_1 \cup A_2)) = \mu(S \cap A_1) + \mu(S \cap A_2)$ which is proved by simply using Definition 1.2 with A_1 in place of A and $S \cap (A_1 \cup A_2)$ in place of S.

Step 3: For each $N = 1, 2, \ldots$, $A_1, A_2, \ldots, A_N \in \mathcal{M} \Rightarrow \cup_{j=1}^{N} A_j \in \mathcal{M}$, which follows from Step 2 by induction on N. Using the additional additivity conclusion of Step 2 we also conclude the additivity $\mu(S \cap (\cup_{j=1}^N A_j)) = \sum_{j=1}^N \mu(S \cap A_j)$ provided $A_1, A_2, \ldots, A_N$ are pairwise disjoint.

Step 4: If $A_1, A_2, \ldots$ are pairwise disjoint sets in $\mathcal{M}$ then $\cup_{j=1}^\infty A_j \in \mathcal{M}$ and furthermore $\mu(S \cap (\cup_{j=1}^\infty A_j)) = \sum_{j=1}^\infty \mu(S \cap A_j)$ for each $S \subset X$. To check this we use the conclusions of Step 3 to observe

$$\begin{aligned}\mu(S) &= \mu(S \cap (\cup_{j=1}^N A_j)) + \mu(S \setminus (\cup_{j=1}^N A_j)) \\ &\geq \mu(S \cap (\cup_{j=1}^N A_j)) + \mu(S \setminus (\cup_{j=1}^\infty A_j)) \\ &= \textstyle\sum_{j=1}^N \mu(S \cap A_j) + \mu(S \setminus (\cup_{j=1}^\infty A_j)).\end{aligned}$$

Since $\sum_{j=1}^N \mu(S \cap A_j) \to \sum_{j=1}^\infty \mu(S \cap A_j) \geq \mu(S \cap (\cup_{j=1}^\infty A_j))$, in view of Remark 1.3 this completes the proof of Step 4, and also establishes the additivity property of Remark 1.7.

Step 5: $A_1, A_2, \ldots \in \mathcal{M} \Rightarrow \cup_{j=1}^\infty A_j \in \mathcal{M}$ (i.e. we do indeed have that $\mathcal{M}$ has property 1.4(3)). To check this, observe that $\cup_{j=1}^\infty A_j = \cup_{j=1}^\infty \widetilde{A}_j$, where $\widetilde{A}_j = A_j \setminus (\cup_{i=0}^{j-1} A_i)$, with $A_0 = \emptyset$. Then $\widetilde{A}_j \in \mathcal{M}$ by Step 1 and Step 3. Since the $\widetilde{A}_j$ are pairwise disjoint we can then apply Step 4 to complete the proof. □

Observe that by 1.7 we have

1.8 $\quad A_j$ μ-measurable, $A_j \subset A_{j+1}\ \forall\, j \geq 1 \Rightarrow \lim_{j\to\infty} \mu(A_j) = \mu(\cup_{j=1}^\infty A_j)$,

because we can write $\cup_{j=1}^\infty A_j = \cup_{j=1}^\infty (A_j \setminus A_{j-1})$ with $A_0 = \emptyset$, and hence, by 1.7,

$$\begin{aligned}\mu(\cup_{j=1}^\infty A_j) &= \textstyle\sum_{j=1}^\infty \mu(A_j \setminus A_{j-1}) = \lim \sum_{j=1}^n \mu(A_j \setminus A_{j-1}) \\ &= \lim \mu(\cup_{j=1}^n (A_j \setminus A_{j-1})) = \lim \mu(A_n),\end{aligned}$$

where at the last step we used $\cup_{j=1}^n (A_j \setminus A_{j-1}) = A_n$.

If $A_1 \supset A_2 \supset \ldots$ then, for each i, $A_i \setminus \cap_{j=1}^\infty A_j = \cup_{j=1}^\infty (A_i \setminus A_j)$, and hence 1.8 implies $\lim_{j\to\infty} \mu(A_i \setminus A_j) = \mu(A_i \setminus \cap_{j=1}^\infty A_j)$, and if $\mu(A_i) < \infty$ this gives $\mu(A_i) - \lim_j \mu(A_j) = \mu(A_i) - \mu(\cap_{j=1}^\infty A_j)$, and hence

1.9 $\quad A_j$ μ-measurable and $A_{j+1} \subset A_j$ for each $j = 1, 2, \ldots$

$$\Rightarrow \lim_{j\to\infty} \mu(A_j) = \mu(\cap_{j=1}^\infty A_j), \text{ provided } \mu(A_i) < \infty \text{ for some } i.$$

An outer measure μ on X is said to be *regular* if for each subset $A \subset X$ there is a μ-measurable subset $B \supset A$ with $\mu(B) = \mu(A)$.

1.10 Remark: If $A_i \subset A_{i+1}\ \forall i$ and μ is regular, then the identity in 1.8 is valid, i.e.

$$\lim_{i\to\infty} \mu(A_i) = \mu(\cup_{i=1}^\infty A_i),$$

even if the A_i are not assumed to be μ-measurable, because for each i we can select a μ-measurable set $\widetilde{A}_i \supset A_i$ with $\mu(\widetilde{A}_i) = \mu(A_i)$, and then $\widehat{A}_i = \cap_{j=i}^\infty \widetilde{A}_j\,(\supset A_i)$ is increasing with $\mu(\widehat{A}_i) = \mu(A_i)$ and $\lim \mu(A_i) \leq \mu(\cup_{i=1}^\infty A_i) \leq \mu(\cup_{i=1}^\infty \widehat{A}_i) = \lim \mu(\widehat{A}_i)$ (by 1.8) $= \lim \mu(A_i)$.

In case X is a topological space, an outer measure μ on X is said to be *Borel-regular* if all Borel sets (see Remark 1.5(2)) are μ-measurable and if for each subset $A \subset X$ there is a Borel set $B \supset A$ such that $\mu(B) = \mu(A)$. (Notice that this does *not* imply $\mu(B \setminus A) = 0$ unless A is μ-measurable and $\mu(A) < \infty$.)

1.11 Remark: There is a close relationship between Borel regular outer measures on a topological space X and abstract Borel measures μ_0 on X. (Recall that a Borel measure μ_0 on X is a map $\mu_0 : \{\text{all Borel sets}\} \to [0, \infty]$ such that (i) $\mu_0(\emptyset) = 0$, and (ii) $\mu_0(\cup_{j=1}^\infty B_j) = \sum_{j=1}^\infty \mu_0(B_j)$ whenever $B_1, B_2, \ldots$ are pairwise disjoint Borel sets in X.) In fact if μ_0 is such a Borel measure on X then

$$\mu(A) = \inf_{B \text{ Borel}, B \supset A} \mu_0(B)$$

defines a Borel regular outer measure on X which agrees with μ_0 on the Borel sets; to check μ-measurability of any Borel set B we just check the inequality in 1.3 by first choosing a Borel set $C \supset S$ with $\mu(C) = \mu(S)$. Conversely, if μ is a Borel regular outer measure on X then the restriction of μ to the Borel sets gives us a Borel measure μ_0 on X.

Given any subset $Y \subset X$ and any outer measure μ on X, we can define a new outer measure $\mu \llcorner Y$ on X by

1.12 $$(\mu \llcorner Y)(Z) = \mu(Y \cap Z), \quad Z \subset X.$$

One readily checks (by using $S \cap Y$ in place of S in Definition 1.2) that all μ-measurable subsets are also $(\mu \llcorner Y)$-measurable (even if Y is *not* μ-measurable). It is also easy to check that $\mu \llcorner Y$ is Borel regular whenever μ is, provided Y is μ-measurable with $\mu(Y) < \infty$. To check this, first use Borel regularity of μ to pick a Borel set B_1 with $B_1 \supset Y$ and $\mu(B_1 \setminus Y) = 0$ and a Borel set $B_2 \supset B_1 \setminus Y$ with $\mu(B_2) = 0$. Then given an arbitrary set $A \subset X$ we have

$$\begin{aligned}A &= (A \cap Y) \cup (A \setminus Y) \subset (A \cap Y) \cup (X \setminus Y) \\ &= (A \cap Y) \cup (X \setminus B_1) \cup (B_1 \setminus Y) \subset (A \cap Y) \cup (X \setminus B_1) \cup B_2.\end{aligned}$$

Finally select a Borel set $B_3 \supset A \cap Y$ with $\mu(B_3) = \mu(A \cap Y)$ and observe that then $A \subset (X \setminus B_1) \cup B_2 \cup B_3$ (which is a Borel set) and $(\mu \llcorner Y)((X \setminus B_1) \cup B_2 \cup B_3) = (\mu \llcorner Y)(A)$.

The following theorem, due to Caratheodory and applicable in case X is a metric space with metric d, is particularly useful. In the statement we use the notation

$$\operatorname{dist}(A,B) = \inf\{d(a,b) : a \in A, b \in B\},$$

interpreted as ∞ if A or B is empty.

1.13 Theorem (Caratheodory's Criterion.) *If X is a metric space with metric d and if μ is an outer measure on X with the property*

$$\mu(A \cup B) = \mu(A) + \mu(B) \text{ for all sets } A, B \subset X \text{ with } \operatorname{dist}(A,B) > 0,$$

then all Borel sets are μ-measurable.

Proof: Since the measurable sets form a σ-algebra, it is enough to prove that all closed sets are μ-measurable (because by definition the Borel sets are the smallest σ-algebra containing all the closed sets), so that by Remark 1.3 we have only to check that

$$(*) \qquad \mu(S) \ge \mu(S \setminus C) + \mu(S \cap C)$$

for all sets $S \subset X$ with $\mu(S) < \infty$ and for all closed sets $C \subset X$.

Let $C_j = \{x \in X : \operatorname{dist}(x, C) \le 1/j\}$. Then $\operatorname{dist}(S \setminus C_j, S \cap C) > 0$, hence

$$\mu(S) \ge \mu((S \setminus C_j) \cup (S \cap C)) = \mu(S \setminus C_j) + \mu(S \cap C),$$

and we will have $(*)$ if we can show $\lim_{j\to\infty} \mu(S \setminus C_j) = \mu(S \setminus C)$. To check this, note that since C is closed we can write

$$S \setminus C = \{x \in X : \operatorname{dist}(x, C) > 0\} = \big(S \setminus C_j\big) \cup \big(\cup_{k=j}^{\infty} R_k\big), \; j = 1, 2, \ldots,$$

where $R_k = \{x \in S : \frac{1}{k+1} < \operatorname{dist}(x, C) \le \frac{1}{k}\}$. But then by subadditivity of μ we have

$$\mu(S \setminus C_j) \le \mu(S \setminus C) \le \mu(S \setminus C_j) + \textstyle\sum_{k=j}^{\infty} \mu(R_k),$$

and hence we will have $\lim_{j\to\infty} \mu(S \setminus C_j) = \mu(S \setminus C)$ as required, provided only that $\sum_{k=1}^{\infty} \mu(R_k) < \infty$.

To check this we note that $\operatorname{dist}(R_i, R_j) > 0$ if $j \ge i+2$, and hence by the hypothesis of the theorem and induction on N we have, for each $N \ge 1$,

$$\textstyle\sum_{k=1}^{N} \mu(R_{2k}) = \mu\big(\cup_{k=1}^{N} R_{2k}\big) \le \mu(S) < \infty$$

and

$$\textstyle\sum_{k=1}^{N} \mu(R_{2k-1}) = \mu\big(\cup_{k=1}^{N} R_{2k-1}\big) \le \mu(S) < \infty. \quad \square$$

We next prove some important regularity properties for Borel regular measures which have a suitable σ-finiteness property:

1.14 Definition: We say a Borel regular measure μ on a topological space X is "open σ-finite" if $X = \cup_j V_j$ where V_j is open in X and $\mu(V_j) < \infty$ for each $j = 1, 2, \ldots$.

Of course μ automatically satisfies such a condition if X is a separable metric space and μ is locally finite (i.e. $x \in X \Rightarrow \mu(B_\rho(x)) < \infty$ for some $\rho > 0$).

1.15 Theorem. *Suppose X is a topological space with the property that every closed subset of X is the countable intersection of open sets (this trivially holds e.g. if X is a metric space), suppose μ is an open σ-finite (as in 1.14 above) Borel-regular measure on X. Then*

$$(1) \qquad \mu(A) = \inf_{U \text{ open}, \, U \supset A} \mu(U)$$

for each subset $A \subset X$, and

$$(2) \qquad \mu(A) = \sup_{C \text{ closed}, \, C \subset A} \mu(C)$$

for each μ-measurable subset $A \subset X$.

1.16 Remark: In case X is a Hausdorff space (so compact sets in X are closed) which is σ-compact (i.e. $X = \cup_j K_j$ with K_j compact for each j), then the conclusion (2) in the above theorem guarantees that

$$\mu(A) = \sup_{K \text{ compact}, \, K \subset A} \mu(K)$$

for each μ-measurable subset $A \subset X$ with $\mu(A) < \infty$, because under the above conditions on X any closed set C can be written as the union of an increasing sequence of compact sets.

Proof of 1.15: We assume first that $\mu(X) < \infty$. First note that in this case (2) can be proved by applying (1) to the complement $X \setminus A$. Also, by Borel regularity of the measure μ, it is enough to prove (1) in case A is a Borel set. Then let

$$\mathcal{A} = \{ \text{ Borel sets } A \subset X : (1) \text{ holds}\}.$$

Trivially $\mathcal{A}$ contains all open sets, and we claim that A is closed under both countable unions and intersections, which we check as follows:

If $A_1, A_2, \ldots \in \mathcal{A}$ then for any given $\varepsilon > 0$ there are open $U_1, U_2, \ldots$ with $U_j \supset A_j$ and $\mu(U_j \setminus A_j) \le 2^{-j}\varepsilon$. Then $(\cup_j U_j) \setminus (\cup_k A_k) = \cup_j (U_j \setminus (\cup_k A_k)) \subset \cup_j (U_j \setminus A_j)$ and $(\cap_j U_j) \setminus (\cap_k A_k) = \cap_j (U_j \setminus (\cap_k A_k)) = \cap_j (\cup_k (U_j \setminus A_k)) \subset \cup_k (U_k \setminus A_k)$ so by subadditivity we have $\mu(\cup_{j=1}^{\infty} U_j \setminus (\cup_k A_k)) < \varepsilon$ and $\lim_{N\to\infty} \mu(\cap_{j=1}^{N} U_j \setminus (\cap_k A_k)) = \mu(\cap_{j=1}^{\infty} U_j \setminus (\cap_k A_k)) < \varepsilon$, so both $\cup_k A_k$ and $\cap_k A_k$ are in $\mathcal{A}$ as claimed.

In particular A must also contain the *closed sets*, because any closed set in X can be written as a countable intersection of open sets. Notice however that at this point it is not clear that $\mathcal{A}$ is a σ-algebra since it is not clear that $\mathcal{A}$ is closed under complementation. For this reason, we let $\widetilde{\mathcal{A}} = \{A \in \mathcal{A} : X \setminus A \in \mathcal{A}\}$, which we claim *is* a σ-algebra, since it clearly has properties (1),(2) of σ-algebra, and if $A_1, A_2, \ldots \in \widetilde{\mathcal{A}}$ then $X \setminus A_1, X \setminus A_2, \ldots \in \mathcal{A}$ and hence $\cup_{j=1}^{\infty} A_j$ and $X \setminus (\cup_{j=1}^{\infty} A_j)(= \cap_{j=1}^{\infty}(X \setminus A_j))$ are both in $\mathcal{A}$ (because $\mathcal{A}$ is closed under countable unions and intersections); thus $\cup_{j=1}^{\infty} A_j \in \widetilde{\mathcal{A}}$ and indeed $\widetilde{\mathcal{A}}$ is a σ-algebra as claimed. Thus $\widetilde{\mathcal{A}}$ is a σ-algebra containing all the closed sets, and hence $\widetilde{\mathcal{A}}$ contains all the Borel sets. Thus $\mathcal{A}$ contains all the Borel sets (so actually we conclude that $\mathcal{A}$ is equal to the collection of all Borel subsets of X) and (1), (2) are both proved in case $\mu(X) < \infty$.

In the case $\mu(X) = \infty$ it still suffices to prove (1) when A is a Borel set. For each $j = 1, 2, \ldots$ apply the previous case $\mu(X) < \infty$ to the measure $\mu \llcorner V_j$, $j = 1, 2, \ldots$. Then for each $\varepsilon > 0$ we can select an open $U_j \supset A$ such that

$$\mu(U_j \cap V_j \setminus A) < \frac{\varepsilon}{2^j},$$

and hence (summing over j)

$$\mu\big(\cup_{j=1}^{\infty}(U_j \cap V_j) \setminus A\big) < \varepsilon.$$

Since $\cup_{j=1}^{\infty}(U_j \cap V_j)$ is open and contains A, this completes the proof of (1).

(2) for the case when $\mu(X) = \infty$ also follows by applying (2) in the finite measure case to the measure $\mu \llcorner V_j$, thus giving, for each $\varepsilon > 0$ and each $j = 1, 2, \ldots$, a closed $C_j \subset A$ with $\mu(A \cap V_j \setminus C_j) < 2^{-j}\varepsilon$. Since $(\cup_{j=1}^{\infty} V_j) \setminus (\cup_{k=1}^{\infty} C_k) = \cup_{j=1}^{\infty}(V_j \setminus (\cup_{k=1}^{\infty} C_k)) \subset \cup_{j=1}^{\infty}(V_j \setminus C_j)$, this gives, by countable subadditivity of μ, $\mu(A \setminus (\cup_{k=1}^{\infty} C_k)) < \varepsilon$. Thus either $\mu(A) = \infty$ and $\mu(\cup_{j=1}^{N} C_j) \to \infty$ or else $\mu(A) < \infty$ and $\mu(A \setminus (\cup_{j=1}^{N} C_j)) < 2\varepsilon$ for sufficiently large N. In either case this completes the proof of (2). □

Using the above theorem, we can now prove Lusin's Theorem:

1.17 Theorem (Lusin's Theorem.) *Let μ be a Borel regular outer measure on a topological space X having the property that every closed subset can be expressed as the countable intersection of open sets (e.g. X is a metric space), let A be any μ-measurable subset of X with $\mu(A) < \infty$, and let $f : A \to \mathbb{R}$ be μ-measurable. Then for each $\varepsilon > 0$ there is a closed set $C \subset X$ with $C \subset A$, $\mu(A \setminus C) < \varepsilon$, and $f|C$ continuous.*

Proof: For each $i = 1, 2, \ldots$ and $j = 0, \pm 1, \pm 2, \ldots$ let

$$A_{ij} = f^{-1}[\tfrac{j-1}{i}, \tfrac{j}{i}),$$

so that A_{ij}, $j = 0, \pm 1, \pm 2, \ldots$, are pairwise disjoint sets in A and $\cup_{j=-\infty}^{\infty} A_{ij} = A$ for each $i = 1, 2, \ldots$. By the remarks following 1.12, we know that $\mu \llcorner A$ is Borel regular, and since it is finite we can apply Theorem 1.15, so for each given $\varepsilon > 0$ there are closed sets C_{ij} in X with $C_{ij} \subset A_{ij}$ and $\mu(A_{ij} \setminus C_{ij}) = (\mu \llcorner A)(A_{ij} \setminus C_{ij}) < 2^{-i-|j|-2}\varepsilon$, hence $\mu(A_{ij} \setminus (\cup_{\ell=-\infty}^{\infty} C_{i\ell})) < 2^{-i-|j|-2}\varepsilon$, hence $\mu(A \setminus (\cup_{j=-\infty}^{\infty} C_{ij})) < 2^{-i}\varepsilon$. So for each $i = 1, 2, \ldots$ there is an integer $J(i)$ with $\mu(A \setminus (\cup_{|j| \le J(i)} C_{ij})) < 2^{-i}\varepsilon$. Since $A \setminus (\cap_{i=1}^{\infty}(\cup_{|j| \le J(i)} C_{ij})) = \cup_{i=1}^{\infty}(A \setminus (\cup_{|j| \le J(i)} C_{ij}))$ this implies $\mu(A \setminus C) < \varepsilon$, where $C = \cap_{i=1}^{\infty}(\cup_{|j| \le J(i)} C_{ij})$, which is a closed set in X.

Finally, define $g_i : \cup_{|j| \le J(i)} C_{ij} \to \mathbb{R}$ by setting $g_i(x) = \frac{j-1}{i}$ on C_{ij}, $|j| \le J(i)$. Then, since the $C_{i1}, \ldots, C_{iJ(i)}$ are pairwise disjoint closed sets, g_i is clearly continuous and its restriction to C is continuous for each i. Furthermore by construction $0 \le f(x) - g_i(x) \le 1/i$ for each $x \in C$ and each $i = 1, 2, \ldots$, so $g_i|C$ converges uniformly to $f|C$ on C, and hence $f|C$ is continuous. □

2 Hausdorff Measure

In this section we suppose X is a metric space with metric d, and we let

$$\omega_m = \frac{\pi^{m/2}}{\Gamma(\frac{m}{2} + 1)}, \quad m \ge 0,$$

where Γ is the Gamma Function $\Gamma(q) = \int_0^{\infty} t^{q-1} e^{-t}\, dt$ for $q > 0$, so that in particular ω_m is the volume (Lebesgue measure) of the unit ball $B_1^m(0)$ in $\mathbb{R}^m$ in case m happens to be a positive integer.

For any $m \ge 0$ we define the m-dimensional Hausdorff (outer) measure

2.1 $$\mathcal{H}^m(A) = \lim_{\delta \downarrow 0} \mathcal{H}^m_\delta(A), \ A \subset X,$$

where, for each $\delta > 0$, $\mathcal{H}^m_\delta(A)$ (called the "size δ approximation to $\mathcal{H}^m$") is defined by taking $\mathcal{H}^m_\delta(\emptyset) = 0$ and, for any non-empty $A \subset X$,

2.2 $$\mathcal{H}^m_\delta(A) = \omega_m \inf \textstyle\sum_{j=1}^{\infty}\Big(\dfrac{\operatorname{diam} C_j}{2}\Big)^m,$$

where the inf is taken over all countable collections $C_1, C_2, \ldots$ of subsets of X such that $\operatorname{diam} C_j < \delta$ and $A \subset \cup_{j=1}^{\infty} C_j$; the right side is to be interpreted as ∞ in case there is no such collection $C_1, C_2, \ldots$. (Of course in a separable metric space X there are always such collections $C_1, C_2, \ldots$ for each $\delta > 0$.) The limit in 2.1 of course always exists (although it may be $+\infty$) because $\mathcal{H}^m_\delta(A)$ is a decreasing function of δ; thus $\mathcal{H}^m(A) = \sup_{\delta > 0} \mathcal{H}^m_\delta(A)$ for each $m \ge 0$. It is left as an exercise to check that $\mathcal{H}^m_\delta$ and $\mathcal{H}^m$ are indeed outer measures on X.

Notice also that $\mathcal{H}^0$ is just "counting measure": $\mathcal{H}^0(\emptyset) = 0$, $\mathcal{H}^0(A) =$ the number of elements in the set A if A is finite, and $\mathcal{H}^0(A) = \infty$ if A is not finite.

2.3 Remarks: (1) Since $\operatorname{diam} C_j = \operatorname{diam} \overline{C}_j$ we can add the additional requirement in the identity 2.2 that the C_j be *closed* without changing the value of $\mathcal{H}^m(A)$; indeed since for any $\varepsilon > 0$ we can find an open set $U_j \supset C_j$ with $\operatorname{diam} U_j < \operatorname{diam} C_j + \varepsilon/2^j$, we could also take the C_j to be *open*.

(2) Evidently $\mathcal{H}^m_\delta(A) < \infty\ \forall m \geq 0$, $\delta > 0$ in case A is a totally bounded subset of the metric space X.

One easily checks from the definition of $\mathcal{H}^m_\delta$ that

2.4 $$\mathcal{H}^m_\delta(A \cup B) = \mathcal{H}^m_\delta(A) + \mathcal{H}^m_\delta(B)\ \forall \delta < \frac{1}{2} d(A, B),$$

hence

2.5 $$\mathcal{H}^m(A \cup B) = \mathcal{H}^m(A) + \mathcal{H}^m(B) \quad \text{whenever} \quad d(A, B) > 0,$$

and therefore all Borel sets are $\mathcal{H}^m$-measurable by the Caratheodory Criterion (Theorem 1.13). It follows from this and Remark 2.3(1) that

2.6 $$\mathcal{H}^m \text{ is Borel-regular for each } m \geq 0.$$

Note: It is *not* true in general that the Borel sets are $\mathcal{H}^m_\delta$-measurable for $\delta > 0$; for instance if $n = 2$ then one easily checks that the half-space $H = \{x = (x^1, x^2) \in \mathbb{R}^n : x^2 > 0\}$ is not $\mathcal{H}^1_\delta$-measurable, because for example it does not cut the set $S_\varepsilon = ([0,1) \times \{0\}) \cup ([0,1) \times \{\varepsilon\})$ additively for sufficiently small ε. Indeed one can directly use the definition of $\mathcal{H}^1_\delta$ to check that $\mathcal{H}^1_\delta(S_\varepsilon) \downarrow 1$ as $\varepsilon \downarrow 0$ (and in particular $\mathcal{H}^1_\delta(S_\varepsilon) < 2$ for sufficiently small ε), whereas $\mathcal{H}^1_\delta(S_\varepsilon \cap H) = \mathcal{H}^1_\delta(S_\varepsilon \setminus H) = 1$.

We will later show that, for each integer $n \geq 1$, $\mathcal{H}^n$ agrees with the usual definition of n-dimensional volume measure on an n-dimensional C^1 submanifold of $\mathbb{R}^{n+k}$, $k \geq 0$. As a first step we want to prove that $\mathcal{H}^n$ and $\mathcal{L}^n$ (n-dimensional Lebesgue measure) agree on $\mathbb{R}^n$.

First recall the standard definition of Lebesgue outer measure $\mathcal{L}^n$ in $\mathbb{R}^n$:

If $\mathcal{K}$ denotes the collection of all n-dimensional intervals I of the form $I = (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)$, where $a_i, b_i \in \mathbb{R}$ and $b_i - a_i > 0$, and if $|I| =$ volume of $I\,(= (b_1 - a_1) \cdots (b_n - a_n))$, then

2.7 $$\mathcal{L}^n(A) = \inf \textstyle\sum_j |I_j|$$

where the inf is taken over all countable (or finite) collections $\{I_1, I_2, \ldots\} \subset \mathcal{K}$ with $A \subset \cup_j I_j$. One easily checks that $\mathcal{L}^n$ is *uniquely characterized* among outer measures on $\mathbb{R}^n$ by the properties

2.8 $$\mathcal{L}^n(I) = |I|\ \forall I \in \mathcal{K}, \quad \mathcal{L}^n(A) = \inf_{U \supset A,\ U \text{ open}} \mathcal{L}^n(U)\ \forall A \subset \mathbb{R}^n.$$

We claim that, on $\mathbb{R}^n$, the outer measures $\mathcal{L}^n, \mathcal{H}^n, \mathcal{H}^n_\delta$ all coincide (for each $\delta > 0$):

2.9 Theorem.

$$\mathcal{L}^n(A) = \mathcal{H}^n(A) = \mathcal{H}^n_\delta(A) \textit{ for every } A \subset \mathbb{R}^n \textit{ and every } \delta > 0.$$

Proof of 2.9: We first show

(1) $$\mathcal{H}^n_\delta(A) \leq \mathcal{L}^n(A)\ \forall \delta > 0$$

as follows: Choose any collection $I_1, I_2, \ldots \in \mathcal{K}$ so that $A \subset \cup_k I_k$. Now for each bounded open set $U \subset \mathbb{R}^n$ and each $\delta > 0$ we can select a pairwise disjoint family of closed balls $B_1, B_2, \ldots$ with $\cup_{j=1}^\infty B_j \subset U$, $\operatorname{diam} B_j < \delta$, and $\mathcal{L}^n(U \setminus \cup_{j=1}^\infty B_j) = 0$. (To see this first decompose U as a union $\cup_{j=1}^\infty C_j$ of closed cubes C_j of diameter $< \delta$ and with pairwise disjoint interiors, and for each $j \geq 1$ select a ball $B_j \subset$ interior C_j with $\operatorname{diam} B_j > \frac{1}{2}$ edge-length of C_j. Then $\mathcal{L}^n(C_j \setminus B_j) < (1 - \theta_n)\mathcal{L}^n(C_j)$, $\theta_n = \omega_n/4^n$, and hence $\mathcal{L}^n(U \setminus \cup_{j=1}^\infty B_j) = \mathcal{L}^n(\cup_{j=1}^\infty (C_j \setminus B_j)) < (1-\theta_n)\mathcal{L}^n(U)$. Thus $\mathcal{L}^n(U \setminus \cup_{j=1}^N B_j) \leq (1-\theta_n)\mathcal{L}^n(U)$ for suitably large N. Since $U \setminus \cup_{j=1}^N B_j$ is open, we can repeat the argument inductively to obtain the required collection of balls.) Then take $U = I_k$ and such a collection of balls $\{B_j\}$. Since $\mathcal{L}^n(Z) = 0 \Rightarrow \mathcal{H}^n_\delta(Z) = 0$ for each subset $Z \subset X$ (by Definitions 2.2, 2.7) we then have (writing $\rho_j =$ radius B_j)

(2) $$\begin{aligned}\mathcal{H}^n_\delta(I_k) &= \mathcal{H}^n_\delta(\cup_{j=1}^\infty B_j) \leq \textstyle\sum_{j=1}^\infty \omega_n \rho_j^n \\ &= \textstyle\sum_{j=1}^\infty \mathcal{L}^n(B_j) = \mathcal{L}^n(\cup_{j=1}^\infty B_j) = \mathcal{L}^n(I_k) = |I_k|,\end{aligned}$$

and hence

(3) $$\mathcal{H}^n_\delta(A) \leq \mathcal{H}^n_\delta(\cup_k I_k) \leq \textstyle\sum_k \mathcal{H}^n_\delta(I_k) \leq \sum_k |I_k|.$$

The proof of (1) is then completed by taking inf over all such collections $\{I_k\}$. □

To prove the reverse inequality we first need a geometric result concerning Lebesgue measure, known as *the isodiametric inequality*:

2.10 Theorem (Isodiametric Inequality.)

$$\mathcal{L}^n(A) \leq \omega_n \left(\frac{\operatorname{diam} A}{2}\right)^n \textit{ for every set } A \subset \mathbb{R}^n.$$

Remark: Thus among all sets $A \subset \mathbb{R}^n$ with a given diameter ρ, the ball with diameter ρ has the largest $\mathcal{L}^n$ measure.

Proof of 2.10: Observe that it suffices to prove this for compact sets because $\bar{A}$ has the same diameter as A and the inequality is trivial if $\operatorname{diam} A = \infty$. For a compact set A we proceed to use *Steiner symmetrization*: The Steiner symmetrization $\mathcal{S}_j(A)$ of the compact set A with respect to the j-th coordinate plane $x^j = 0$ is defined as follows: For ξ in the coordinate plane $x^j = 0$ let $\ell_j = \{\xi + te_j : t \in \mathbb{R}\}$ and let π is the projection $\xi + te_j \mapsto t$, taking the line $\ell_j(\xi) = \{\xi + te_j : t \in \mathbb{R}\}$ onto the real line $\mathbb{R}$, and let $\sigma_j(A, \xi)$ denote the closed line segment $\{\xi + te_j : |t| \le \frac{1}{2}\mathcal{L}^1(\pi(A \cap \ell_j(\xi)))\}$. Then

$$\mathcal{S}_j(A) = \cup_{\{\xi : A\cap\ell_j(\xi)\neq\emptyset\}}\ \sigma_j(A, \xi).$$

(Thus $\mathcal{S}_j(A)$ is obtained by replacing $A \cap \ell_j(\xi)$ with the segment $\sigma_j(A, \xi)$ for each ξ such that $A \cap \ell_j(\xi) \neq \emptyset$.) This process gives a new compact set (see the note below) $\mathcal{S}_j(A)$ with the diameter not larger than the diameter of the original set A and, by Fubini, the same Lebesgue measure. Further if $i \neq j$ and if A is already invariant under reflection in the i-th coordinate plane $x^i = 0$, then by definition $\mathcal{S}_j(A)$ is invariant under reflection in both the i-th and the j-th coordinate planes. Thus by applying Steiner symmetrization successively with respect to coordinate planes $x^1 = 0, x^2 = 0, \ldots, x^n = 0$, we get a new compact set $\widetilde{A}$ with diameter $\le \operatorname{diam} A$, having the same Lebesgue measure as A, and being invariant with respect to the transformation $x \mapsto -x$. In particular this means that $\widetilde{A}$ is contained in the closed ball with radius $\frac{1}{2} \operatorname{diam} A$ and center 0, whence

$$\mathcal{L}^n(A) = \mathcal{L}^n(\widetilde{A}) \le \omega_n(\frac{1}{2} \operatorname{diam} A)^n$$

as required. □

Note: In the above we used the fact that $\mathcal{S}_j(A)$ is compact if A is compact. That is clearly true because $\mathcal{L}^1(\pi(A \cap \ell_j(\xi)))$ (where $\ell_j(\xi), \pi$ are as in the above proof) is an upper semi-continuous function of ξ if ξ is restricted to lie in the j-th coordinate hyperplane $x^j = 0$. To check this upper semi-continuity, observe that if $\varepsilon > 0$ we can select open $U \subset \mathbb{R}$ with $\pi(A \cap \ell_j(\xi)) \subset U$ and $\mathcal{L}^1(U) \le \mathcal{L}^1(\pi(A \cap \ell_j(\xi))) + \varepsilon$. Since A is closed and $\pi^{-1}(U)$ is an open set containing the compact set $A \cap \ell_j(\xi)$, we then see that for any sequence $\xi_i \to \xi$ in the plane $x^j = 0$ we must have $A \cap \ell_j(\xi_i) \subset \pi^{-1}(U)$ for sufficiently large i, and hence $\pi(A \cap \ell_j(\xi_i)) \subset U$ for sufficiently large i, thus giving $\mathcal{L}^1(\pi(A \cap \ell_j(\xi_i))) \le \mathcal{L}^1(\pi(A \cap \ell_j(\xi))) + \varepsilon$ for all sufficiently large i.

Completion of the proof of 2.9: We have to prove

$$(\ddagger) \qquad \mathcal{L}^n(A) \le \mathcal{H}^n_\delta(A) \ \ \forall \delta > 0,\ A \subset \mathbb{R}^n.$$

Suppose $\delta > 0$, $A \subset \mathbb{R}^n$, and let $C_1, C_2, \ldots$ be any countable collection with $A \subset \cup_j C_j$ and $\operatorname{diam} C_j < \delta$. Then

$$\begin{aligned}\mathcal{L}^n(A) &\le \mathcal{L}^n(\cup_j C_j) \le \textstyle\sum_j \mathcal{L}^n(C_j)\\ &\le \textstyle\sum_j \omega_n \left(\frac{\operatorname{diam} C_j}{2}\right)^n \text{ by 2.10.}\end{aligned}$$

Taking the inf over all such collections $\{C_j\}$ we have $(\ddagger)$ as required. □

3 Densities

Throughout this section X will denote a metric space with metric d. We first we want to introduce the notion of n-dimensional density of a measure μ on X, where X continues to denote a metric space with metric d. For any outer measure μ on X, any subset $A \subset X$, and any point $x \in X$, we define the n-dimensional upper and lower densities $\Theta^{*n}(\mu, A, x)$, $\Theta^n_*(\mu, A, x)$ by

3.1

$$\begin{aligned}\Theta^{*n}(\mu, A, x) &= \limsup_{\rho\downarrow 0} \frac{\mu(A \cap B_\rho(x))}{\omega_n \rho^n}\\ \Theta^n_*(\mu, A, x) &= \liminf_{\rho\downarrow 0} \frac{\mu(A \cap B_\rho(x))}{\omega_n \rho^n}.\end{aligned}$$

In case $A = X$ we simply write $\Theta^{*n}(\mu, x)$ and $\Theta^n_*(\mu, x)$ to denote these quantities so that $\Theta^{*n}(\mu, A, x) = \Theta^{*n}(\mu \llcorner A, x)$, $\Theta^n_*(\mu, A, x) = \Theta^n_*(\mu \llcorner A, x)$.

3.2 Remark: One readily checks that if all Borel sets are μ-measurable and $\mu(B_\rho(x))$ is finite on each ball $B_\rho(x) \subset X$, then $\mu(A \cap B_\rho(x)) \ge \limsup_{y\to x} \mu(A \cap B_\rho(y))$ for each fixed $\rho > 0$ (i.e. $(\omega_n\rho^n)^{-1}\mu(A \cap B_\rho(x))$ is an upper semi-continuous function of x for each ρ), hence $\inf_{0<\rho<\delta}(\omega_n\rho^n)^{-1}\mu(A \cap B_\rho(x))$ is also upper semi-continuous and hence Borel measurable. Thus

$$\begin{aligned}\Theta^n_*(\mu, A, x) &= \lim_{\delta\downarrow 0}\ \inf_{0<\rho<\delta}\ (\omega_n\rho^n)^{-1}\,\mu(A \cap B_\rho(x))\\ &= \lim_{j\to\infty}\ \inf_{0<\rho<1/j}(\omega_n\rho^n)^{-1}\mu(A \cap B_\rho(x))\end{aligned}$$

is also Borel measurable. Similarly since $\mu(A \cap \breve{B}_\rho(x))$ is lower semi-continuous (where $\breve{B}_\rho(x)$ denotes the open ball of radius ρ and center x) and $\Theta^{*n}(\mu, A, x)$ can be written $\lim_{j\to\infty} \sup_{0<\rho<1/j}(\omega_n\rho^n)^{-1}\mu(A \cap \breve{B}_\rho(x))$, we also have $\Theta^{*n}(\mu, A, x)$ is Borel measurable.

Subsequently we use the notation that if $\Theta^{*n}(\mu, A, x) = \Theta^n_*(\mu, A, x)$ then the common value will be denoted $\Theta^n(\mu, A, x)$.

Appropriate information about the upper density gives connections between μ and $\mathcal{H}^n$. Specifically, we have the following comparison theorem:

3.3 Theorem. *Let μ be a Borel-regular measure on the metric space X, $t \geq 0$, and $A_1 \subset A_2 \subset X$. Then*

$$\Theta^{*n}(\mu, A_2, x) \geq t \ \forall\, x \in A_1 \Rightarrow t\mathcal{H}^n(A_1) \leq \mu(A_2),$$

An important special case of this theorem is the case $A_1 = A_2$. Notice that we do *not* need to assume A_1, A_2 are μ-measurable here.

The proof of 3.3 will make use of the following important "5-times covering lemma," in which we use the notation that if B is a ball $B_\rho(x) \subset X$, then $\widehat{B} = B_{5\rho}(x)$.

3.4 Lemma (5-times Covering Lemma). *If $\mathcal{B}$ is any family of closed balls in X with $R \equiv \sup\{\operatorname{diam} B : B \in \mathcal{B}\} < \infty$, then there is a pairwise disjoint subcollection $\mathcal{B}' \subset \mathcal{B}$ such that*

$$\cup_{B \in \mathcal{B}} B \subset \cup_{B \in \mathcal{B}'} \widehat{B};$$

in fact we can arrange the stronger property

$$(\ddagger)\qquad B \in \mathcal{B} \Rightarrow \exists\, B' \in \mathcal{B}' \text{ with } B' \cap B \neq \emptyset \text{ and } \operatorname{diam}(B') \geq \frac{1}{2}\operatorname{diam}(B), \text{ hence } \widehat{B'} \supset B.$$

Proof: For $j = 1, 2, \ldots$ let $\mathcal{B}_j = \{B \in \mathcal{B} : R/2^j < \operatorname{diam} B \leq R/2^{j-1}\}$, so that $\mathcal{B} = \cup_{j=1}^\infty \mathcal{B}_j$, and this is a pairwise disjoint union. Proceed to define $\mathcal{B}'_j \subset \mathcal{B}_j$ as follows:

(i) Let $\mathcal{B}'_1$ be any maximal pairwise disjoint subcollection of $\mathcal{B}_1$. Such $\mathcal{B}'_1$ exist by applying Zorn's lemma to $\mathcal{C} = \{\mathcal{A} : \mathcal{A} \text{ is a pairwise disjoint subcollection of } \mathcal{B}_1\}$, which is partially ordered by inclusion; notice for any totally ordered subcollection $\mathcal{T} \subset \mathcal{C}$ we clearly have $\cup_{\mathcal{A} \in \mathcal{T}} \mathcal{A} \in \mathcal{C}$, so Zorn's lemma is indeed applicable. Notice also that in a general metric space the collection $\mathcal{B}'_1$ could be uncountable, but of course in a separable metric space (i.e. a metric space with a countable dense subset) all pairwise disjoint collections of balls must be countable.

(ii) Assuming $j \geq 2$ and that $\mathcal{B}'_1 \subset \mathcal{B}_1, \ldots, \mathcal{B}'_{j-1} \subset \mathcal{B}_{j-1}$ are defined, let $\mathcal{B}'_j$ be a maximal pairwise disjoint subcollection of $\{B \in \mathcal{B}_j : B \cap B' = \emptyset \text{ whenever } B' \in \cup_{i=1}^{j-1} \mathcal{B}'_i\}$. Again, Zorn's lemma guarantees such a maximal collection exists.

Now if $j \geq 1$ and $B \in \mathcal{B}_j$ we must have

$$B \cap B' \neq \emptyset \text{ for some } B' \in \cup_{i=1}^j \mathcal{B}'_i$$

(otherwise we contradict maximality of $\mathcal{B}'_j$), and for such a pair B, B' we have $\operatorname{diam} B \leq R/2^{j-1} = 2R/2^j \leq 2 \operatorname{diam} B'$, so that $(\ddagger)$ holds; in particular $B \subset \widehat{B'}$. □

In the following corollary we use the terminology that a subset $A \subset X$ is *covered finely* by a collection $\mathcal{B}$ of balls, meaning that for each $x \in A$ and each $\varepsilon > 0$, there is a ball $B \in \mathcal{B}$ with $x \in B$ and $\operatorname{diam} B < \varepsilon$.

3.5 Corollary. *If $\mathcal{B}$ is as in* Theorem 3.4, *if A is a subset of X covered finely by $\mathcal{B}$, and if $\mathcal{B}' \subset \mathcal{B}$ is any pairwise disjoint subcollection of $\mathcal{B}$ satisfying* 3.4 $(\ddagger)$, *then*

$$A \setminus \cup_{j=1}^N B_j \subset \cup_{B \in \mathcal{B}' \setminus \{B_1, \ldots, B_N\}} \widehat{B}$$

for each finite subcollection $\{B_1, \ldots, B_N\} \subset \mathcal{B}'$.

Proof: Let $x \in A \setminus \cup_{j=1}^N B_j$. Since $\mathcal{B}$ covers A finely and since $X \setminus \cup_{j=1}^N B_j$ is open, we can then find $B \in \mathcal{B}$ with $B \cap (\cup_{j=1}^N B_j) = \emptyset$ and $x \in B$, and (by 3.4$(\ddagger)$) find $B' \in \mathcal{B}'$ with $B' \cap B \neq \emptyset$ and $\widehat{B'} \supset B$. Evidently $B' \neq B_j \ \forall\, j = 1, \ldots, N$, so $x \in \cup_{B' \in \mathcal{B}' \setminus \{B_1, \ldots, B_N\}} \widehat{B'}$. □

Proof of 3.3: We can assume $\mu(A_2) < \infty$ and $t > 0$ otherwise the result is trivial. Take $\tau \in (0, t)$, so that then

$$\Theta^{*n}(\mu, A_2, x) > \tau \text{ for } x \in A_1.$$

For $\delta > 0$, let $\mathcal{B}$ (depending on δ) be defined by

$$\mathcal{B} = \{\text{closed balls } B_\rho(x) : x \in A_1,\ 0 < \rho < \delta/2,\ \mu(A_2 \cap B_\rho(x)) \geq \tau\omega_n\rho^n\}.$$

Evidently $\mathcal{B}$ covers A_1 finely and hence there is a disjoint subcollection $\mathcal{B}' \subset \mathcal{B}$ so that 3.4 $(\ddagger)$ holds. Since $\mu(A_2 \cap B) > 0$ for each $B \in \mathcal{B}$ and since $B_1, \ldots, B_N \in \mathcal{B}' \Rightarrow \sum_{j=1}^N \mu(A_2 \cap B_j) = \mu(A_2 \cap (\cup_{j=1}^N B_j)) \leq \mu(A_2) < \infty$ it follows that $\mathcal{B}'$ is a countable collection $\{B_{\rho_1}(x_1), B_{\rho_2}(x_2), \ldots\}$ and hence 3.5 implies

$$A_1 \setminus \cup_{j=1}^N B_{\rho_j}(x_j) \subset \cup_{j=N+1}^\infty B_{5\rho_j}(x_j) \ \forall N \geq 1.$$

and also $\tau \sum_{j=1}^\infty \omega_n \rho_j^n \leq \sum_{j=1}^\infty \mu(A_2 \cap B_{\rho_j}(x_j)) = \mu(A_2 \cap (\cup_{j=1}^\infty B_{\rho_j}(x_j))) \leq \mu(A_2) < \infty$. Thus $A_2 \subset (\cup_{j=1}^N B_{\rho_j}(x_j)) \cup (\cup_{j=N+1}^\infty B_{5\rho_j}(x_j))$ and hence by the definition 2.2 of $\mathcal{H}^n_\delta$ we have

$$\mathcal{H}^n_{5\delta}(A_1) \leq \sum_{j=1}^N \omega_n \rho_j^n + 5^n \sum_{j=N+1}^\infty \omega_n \rho_j^n.$$

Hence letting $N \to \infty$ we deduce

$$\tau\mathcal{H}^n_{5\delta}(A_1) \leq \mu(A_2).$$

Letting $\delta \downarrow 0$ and then $\tau \uparrow t$, we then have the required result. □

As a corollary to 3.3 we can prove the following

3.6 Theorem (Upper Density Theorem.) *If μ is Borel regular on X and if A is a μ-measurable subset of X with $\mu(A) < \infty$, then*

$$\Theta^{*n}(\mu, A, x) = 0 \text{ for } \mathcal{H}^n\text{-a.e. } x \in X \setminus A.$$

3.7 Remarks: (1) Of course $\mu = \mathcal{H}^n$ is an important special case.

(2) If $X = \cup_{j=1}^{\infty} V_j$ with V_j open and $\mu(V_j) < \infty$ for each $j = 1, 2, \ldots$, then one can drop the hypothesis that $\mu(A) < \infty$, because we can apply the theorem with $\mu \llcorner V_j$ in place of μ to conclude that

$$\Theta^{*n}(\mu, A, x) = \Theta^{*n}(\mu, A \cap V_j, x) = 0 \text{ for } \mathcal{H}^n\text{-a.e. } x \in V_j \setminus A, \ j = 1, 2, \ldots,$$

and hence $\Theta^{*n}(\mu, A, x) = 0$ for $\mathcal{H}^n$-a.e. $x \in X \setminus A$.

Proof of 3.6: Let $t > 0$ and $S_t = \{x \in X \setminus A : \Theta^{*n}(\mu, A, x) \geq t\}$, and let C be any closed subset of A. Since $X \setminus C$ is open and $S_t \subset X \setminus A \subset X \setminus C$ we have $\Theta^{*n}(\mu, A \cap (X \setminus C), x) = \Theta^{*n}(\mu, A, x) \geq t$ for $x \in S_t$. Thus we can apply 3.3 with $\mu \llcorner A$, S_t, $X \setminus C$ in place of μ, A_1, A_2 to give $t\mathcal{H}^n(S_t) \leq \mu(A \setminus C)$ for each closed set $C \subset A$. But according to 1.15(2) we can choose closed $C = C_j \subset A$ with $\mu(A \setminus C_j) \to 0$, so we have $\mathcal{H}^n(S_t) = 0$. Taking $t = 1/i$, $i = 1, 2, \ldots$, we thus conclude $\mathcal{H}^n(\{x \in X \setminus A : \Theta^{*n}(\mu, A, x) > 0\}) = 0$. $\square$

Notice that we have the following important corollary to the above theorem:

3.8 Corollary. *If $A \subset \mathbb{R}^n$ is $\mathcal{L}^n$-measurable then the density $\Theta^n(\mathcal{L}^n, A, x)$ exists $\mathcal{L}^n$-a.e. on $\mathbb{R}^n$, and $\Theta^n(\mathcal{L}^n, A, x) = 0$ $\mathcal{L}^n$-a.e. on $\mathbb{R}^n \setminus A$ and $= 1$ $\mathcal{L}^n$-a.e. on A.*

Proof: Indeed $(\omega_n \rho^n)^{-1}\mathcal{L}^n(A \cap B_\rho(x)) + (\omega_n \rho^n)^{-1}\mathcal{L}^n(B_\rho(x) \setminus A) = 1$ for each $\rho > 0$, and, by the Upper Density Theorem, the first term on the left $\to 0$ as $\rho \downarrow 0$ for $\mathcal{L}^n$-a.e. $x \in \mathbb{R}^n \setminus A$ while the second term on the left $\to 0$ as $\rho \downarrow 0$ for $\mathcal{L}^n$-a.e. $x \in A$. $\square$

We next want to discuss the possibility of extending the Comparison and Upper Density Theorems 3.3, 3.6 (and hence the above corollary) to the situation when we consider the upper density of a Borel regular measure μ with respect to another Borel regular measure μ_0. In this case we always assume μ_0 is locally finite, so that for each $x \in X$ there is $\rho > 0$ with $\mu_0(B_\rho(x)) < \infty$. We note that this is automatic if μ_0 is open σ-finite as in 1.14. Then the upper density $\Theta^{*\mu_0}(\mu, x)$ of μ with respect to μ_0 at the point $x \in X$ is defined by

3.9
$$\Theta^{*\mu_0}(\mu, x) = \begin{cases} \limsup_{\rho \downarrow 0} \dfrac{\mu(B_\rho(x))}{\mu_0(B_\rho(x))} & \text{for } x \in X \setminus (U_0 \cup V_0) \\ \infty & \text{for } x \in U_0 \setminus V_0 \\ 0 & \text{for } x \in V_0, \end{cases}$$

where U_0 is the open set consisting of all points $x \in X$ with $\mu_0(B_\rho(x)) = 0$ for some $\rho > 0$ and V_0 is the open set consisting of all points $x \in X$ with $\mu(B_\rho(x)) = 0$ for some $\rho > 0$. Notice $\Theta^{*\mu_0}(\mu, x) = \Theta^{*n}(\mu, x)$ in the special case when $X = \mathbb{R}^n$ and $\mu_0 = \mathcal{L}^n$.

To prove a useful analogue to the Upper Density Theorem 3.6 in this situation we need to assume that μ_0 has the "Symmetric Vitali" property according to the following definition:

3.10 Definition (Symmetric Vitali Property): An outer measure μ_0 on the metric space X is said to have the Symmetric Vitali Property if, given any $A \subset X$ with $\mu_0(A) < \infty$ and and any collection $\mathcal{B}$ of closed balls with centers in A which cover A finely (i.e. $\inf\{\rho : B_\rho(x) \in \mathcal{B}\} = 0$ for each $x \in A$), there is a countable pairwise disjoint collection $\mathcal{B}' = \{B_{\rho_j}(x_j) : j = 1, 2, \ldots\} \subset \mathcal{B}$ with $\mu_0(A \setminus (\cup_{j=1}^{\infty} B_{\rho_j}(x_j))) = 0$.

Before proceeding, we make some important notes concerning the open set U_0 in 3.9 and the Symmetric Vitali Property:

3.11 Remarks: (1) First note that there are various scenarios which guarantee that the open set U_0 in the definition 3.9 of the density $\Theta^{*\mu_0}(\mu, x)$ has μ_0-measure zero. For example if X is separable (i.e. X has a countable dense subset) then any open set, including U_0, can be written as a countable union of balls $B_\rho(x) \subset U_0$, and hence U_0 certainly has μ_0-measure zero in this case. Also, if μ_0 is σ-finite and has the Symmetric Vitali Property, then (because U_0 is trivially covered finely by the balls $B_\rho(x) \subset U_0$) then there is a countable collection of balls contained in U_0 covering μ_0-almost all of U_0), so again $\mu_0(U_0) = 0$.

(2) Observe also that in case X is a separable this Symmetric Vitali Property is satisfied by any Borel regular measure μ_0 with $\mu_0(X) < \infty$ which has the "doubling property" that there is a fixed constant C such that

(‡) $$\mu_0(B_{2\rho}(x)) \leq C\mu_0(B_\rho(x)) \ \forall \text{ closed ball } B_\rho(x) \subset X.$$

Indeed in this case, given $A \subset X$ with $\mu(A) < \infty$ and a collection $\mathcal{B}$ of closed balls which cover A finely, by the Corollary 3.5 of the 5-times Covering Lemma we can select a pairwise disjoint subcollection $\mathcal{B}'$ (which is countable by the separability of

X, hence expressible $\mathcal{B}' = \{B_{\rho_j}(x_j) : j = 1, 2, \ldots\}$) with

$$A \setminus \{B_{\rho_1}(x_1), \ldots, B_{\rho_N}(x_N)\} \subset \cup_{j=N+1}^{\infty} B_{5\rho_j}(x_j)$$

and hence, since $\mu_0(B_{5\rho}(x)) \le \mu_0(B_{8\rho}(x)) \le C^3 \mu_0(B_\rho(x))$ by (‡),

$$\mu_0(A \setminus \{B_{\rho_1}(x_1), \ldots, B_{\rho_N}(x_N)\}) \le C^3 \textstyle\sum_{j=N+1}^{\infty} \mu_0(B_{\rho_j}(x_j)) \to 0 \text{ as } N \to \infty$$

because $\sum_j \mu_0(B_{\rho_j}(x_j)) = \mu_0(\cup_j B_{\rho_j}(x_j)) < \infty$. Thus

$$\mu_0(A \setminus (\cup_{j=1}^{\infty} B_{\rho_j}(x_j))) = 0,$$

as claimed.

(3) A very important fact is that *any* Borel regular measure μ_0 on $\mathbb{R}^n$ which is finite on each compact subset automatically has the Symmetric Vitali Property. In order to check this we'll need the following famous covering lemma due to Besicovitch:

3.12 Lemma (Besicovitch Covering Lemma.) *Suppose $\mathcal{B}$ is a collection of closed balls in $\mathbb{R}^n$, let A be the set of centers, and suppose the set of all radii of balls in $\mathcal{B}$ is a bounded set. Then there are sub-collections $\mathcal{B}_1, \ldots, \mathcal{B}_N \subset \mathcal{B}$ ($N = N(n)$) such that each $\mathcal{B}_j$ is a pairwise disjoint (or empty) collection, and $\cup_{j=1}^N \mathcal{B}_j$ still covers A : $A \subset \cup_{j=1}^N (\cup_{B \in \mathcal{B}_j} B)$.*

We emphasize that N is a certain fixed constant depending only on n. For the proof of this lemma we refer for example to [EG92] or [Fed69].

Proof of Remark 3.11(3): Let μ be a Radon measure on $\mathbb{R}^n$, $A \subset \mathbb{R}^n$ with $\mu(A) < \infty$, $\mathcal{B}$ a collection of closed balls with centers in A covering A finely. By the Besicovitch lemma we can choose $\mathcal{B}_1, \ldots, \mathcal{B}_N \subset \{B \in \mathcal{B} : \operatorname{diam} B \le 1\}$ such that $\cup_{j=1}^N \mathcal{B}_j$ covers A. Then for at least one $j \in \{1, \ldots, N\}$ we get

$$\mu(A \setminus \cup_{B \in \mathcal{B}_j} B) \le (1 - \frac{1}{N})\mu(A)$$

and hence taking a suitable finite subcollection $\{B_1, \ldots, B_Q\} \subset \mathcal{B}_j$,

$$\mu(A \setminus \cup_{k=1}^Q B_k) \le (1 - \frac{1}{2N})\mu(A).$$

Since $\mathcal{B}$ covers A finely, and since $\cup_{k=1}^Q B_k$ is closed, the collection $\widetilde{\mathcal{B}} = \{B \in \mathcal{B} : B \cap (\cup_{k=1}^Q B_k) = \emptyset\}$ covers $A \setminus \cup_{k=1}^Q B_k$ finely, so with $A \setminus \cup_{k=1}^Q B_k$ in place of A the same argument says that we can select new balls $B_{Q+1}, \ldots, B_P \in \widetilde{\mathcal{B}}$ such that

$$\begin{aligned} \mu(A \setminus \cup_{k=1}^P B_k) &\le (1 - \frac{1}{2N})\mu(A \setminus \cup_{k=1}^Q B_k) \\ &\le (1 - \frac{1}{2N})^2 \mu(A). \end{aligned}$$

Continuing (inductively) in this way, we conclude that there is a pairwise disjoint sequence $B_1, B_2, \ldots$ of balls in $\mathcal{B}$ such that

$$\mu(A \setminus \cup_{k=1}^{\infty} B_k) = 0.$$

Thus Remark 3.11(3) is established. □

We now want to prove an analogue of the Comparison Theorem 3.3 in case we use $\Theta^{*\mu_0}(\mu, x)$ of 3.9 in place of the upper density $\Theta^{*n}(\mu, x)$.

3.13 Theorem. *Suppose μ, μ_0 are open σ-finite (as in* 1.14*) Borel regular measures on X, μ_0 and has the Symmetric Vitali Property, and $A \subset X$, $t \ge 0$. Then*

$$\Theta^{*\mu_0}(\mu, x) \ge t \text{ for all } x \in A \Rightarrow \mu(A) \ge t\mu_0(A).$$

Note: A is not assumed to be measurable.

Proof: The proof is similar to the proof of Theorem 3.3, except that we use the Symmetric Vitali Property for μ_0 in place of the 5 times Covering Lemma: First let U_0 be the open set of μ_0 measure zero as in the definition 3.9. As observed in Remark 3.11(1) we have $\mu_0(U_0) = 0$. We can assume without loss of generality that $t > 0$. Let $U \supset A$ be open, $\tau \in (0, t)$, and consider the collection $\mathcal{B}$ of all closed balls $B_\rho(x) \subset U$ with $x \in A \cap X \setminus U_0$ such that $\mu(\cap B_\rho(x)) > \tau\mu_0(B_\rho(x))$. Evidently $\mathcal{B}$ covers $A \cap (X \setminus U_0)$ finely, so by the Symmetric Vitali Property for μ_0 there is a countable pairwise disjoint subcollection $B_{\rho_j}(x_j)$, $j = 1, 2, \ldots$, of $\mathcal{B}$ with $\mu_0(A \setminus (\cup_j B_{\rho_j}(x_j))) = 0$. Then $\mu(A \cap B_{\rho_j}(x_j)) \ge \tau\mu_0(B_{\rho_j}(x_j))$ for each j, and hence by summing we obtain

$$\tau\mu_0(A) \le \mu(\cup_j B_{\rho_j}(x_j)) \le \mu(U).$$

Since $\mu(A) = \inf_{U \text{ open}, U \supset A} \mu(U)$ by Theorem 1.15, we thus have the stated result by letting $\tau \uparrow t$. □

Observe that in particular the above comparison lemma gives

3.14 Corollary. *If μ, μ_0 are as in* Theorem 3.13 *above then $\Theta^{*\mu_0}(\mu, x) < \infty$ for μ_0-a.e. $x \in X$.*

Proof: We are given open V_j with $X = \cup_j V_j$ and $\mu_0(V_j) < \infty$ for each j. Let $\mu_j = \mu \llcorner V_j$, $j = 1, 2, \ldots$. Theorem 3.13, with $A_t = \{x \in V_j \setminus U_0 : \Theta^{*\mu_0}(\mu, x) \ge t\}$ in place of A and μ_j in place of μ, implies

$$t\mu_0(A_t) \le \mu_j(A_t) \le \mu(V_j) \ \forall\, t > 0,$$

so $\mu_0(\{x \in V_j : \Theta^{*\mu_0}(\mu, x) = \infty\}) \le t^{-1}\mu(V_j)$ for each $t > 0$, hence $\mu_0(\{x \in V_j : \Theta^{*\mu_0}(\mu, x) = \infty\}) = 0$ for each j □

As a second corollary we state the following general Upper Density Theorem:

3.15 Theorem (General Upper Density Th.) *If μ, μ_0 are Borel regular on X, if μ_0 open σ-finite (as in* 1.14*) and has the Symmetric Vitali Property, and if A is a μ-measurable subset of X with $\mu(A) < \infty$, then*

$$\Theta^{*\mu_0}(\mu \llcorner A, x) = 0 \text{ for } \mu_0\text{-a.e. } x \in X \setminus A.$$

Proof: XXX The proof is essentially the same as the proof of Theorem 3.6, except that we use the general comparison theorem 3.13 in place of 3.3. The details are left as an exercise. □

Using the above theorem we can now prove the general density theorem:

3.16 Theorem. *If μ is open σ-finite (as in* 1.14*) Borel regular measure on X, if μ has the Symmetric Vitali Property, and if A is a μ-measurable subset of X, then*

$$\lim_{\rho\downarrow 0} \frac{\mu(A \cap B_\rho(x))}{\mu(B_\rho(x))} = \begin{cases} 1 & \mu\text{-a.e. } x \in A \\ 0 & \mu\text{-a.e. } x \in X \setminus A. \end{cases}$$

Proof: Since $X = \cup_j V_j$ with V_j open and $\mu(V_j) < \infty$ for each j, we can assume without loss of generality that $\mu(X) < \infty$. As in Remark 3.11(1) we see that the set of $x \in X$ such that $\mu(B_\rho(x)) = 0$ for some $\rho > 0$ is an open set U_0 with $\mu(U_0) = 0$. For $x \in X \setminus U_0$ we have

$$\frac{\mu(A \cap B_\rho(x))}{\mu(B_\rho(x))} + \frac{\mu(B_\rho(x) \setminus A)}{\mu(B_\rho(x))} = 1 \text{ for each } \rho > 0,$$

and the first term on the left → 0 for μ-a.e. $x \in X \setminus A$ by the Upper Density Theorem 3.15 with $\mu_0 = \mu$, whereas the second term on the left → 0 for μ-a.e. $x \in A$ by the same theorem with $\mu_0 = \mu$ and $X \setminus A$ in place of A. □

The following Lebesgue differentiation theorem is an easy corollary:

3.17 Corollary. *If X, μ are as in* Theorem 3.16 *and if $f : X \to \mathbb{R}$ is locally μ-integrable on X (i.e. f is μ-measurable and $x \in X \Rightarrow \int_{B_\rho(x)} |f|\, d\mu < \infty$ for some $\rho > 0$), then*

$$\lim_{\rho\downarrow 0}(\mu(B_\rho(x)))^{-1}\int_{B_\rho(x)} f\, d\mu = f(x) \text{ for } \mu\text{-a.e. } x \in X$$

Proof: Since $f = \max\{f, 0\} - \max\{-f, 0\}$ we can assume without loss of generality that $f \geq 0$. Let $\nu_0(A) = \int_A f\, d\mu$ for each Borel set $A \subset X$, and let ν be the associated Borel regular outer measure on X as described in Remark 1.11. For each $i, j = 1, 2, \ldots$ let $A_{ij} = f^{-1}[(j-1)/i, j/i)$ and observe that

$$(*) \quad \begin{aligned} \mu(B_\rho(x) \cap A_{ij})(j-1)/i &\leq \int_{A_{ij}\cap B_\rho(x)} f\, d\mu = \nu(A_{ij} \cap B_\rho(x)) \\ &= \int_{B_\rho(x)} f\, d\mu - \nu(B_\rho(x) \setminus A_{ij}) \leq \mu(B_\rho(x) \cap A_{ij})\, j/i. \end{aligned}$$

By Theorem 3.16,

$$\lim_{\rho\downarrow 0}(\mu(B_\rho(x)))^{-1}\mu(B_\rho(x) \cap A_{ij}) = 1 \text{ and } \lim_{\rho\downarrow 0}(\mu(B_\rho(x)))^{-1}\mu(B_\rho(x) \setminus A_{ij})) = 0$$

for μ-a.e. $x \in A_{ij}$, and by the Upper Density Theorem 3.15 (with ν, μ in place of μ, μ_0 respectively), we also have

$$\lim_{\rho\downarrow 0}(\mu(B_\rho(x)))^{-1}\nu(B_\rho(x) \setminus A_{ij}) = 0$$

for μ-a.e. $x \in A_{ij}$. Since $(j-1)/i \leq f(x) \leq j/i$ on A_{ij}, $(*)$ then implies

$$\begin{aligned} f(x) - 1/i &\leq \liminf_{\rho\downarrow 0}(\mu(B_\rho(x)))^{-1}\int_{B_\rho(x)} f\, d\mu \\ &\leq \limsup_{\rho\downarrow 0}(\mu(B_\rho(x)))^{-1}\int_{B_\rho(x)} f\, d\mu \leq f(x) + 1/i \end{aligned}$$

for μ-a.e. $x \in X$, $i = 1, 2, \ldots$. □

Of course we can also take the lower density $\Theta_*^{\mu_0}(\mu, x)$ of μ with respect to μ_0 which we define analogously to the upper density in 3.9, by

3.18
$$\Theta_*^{\mu_0}(\mu, x) = \begin{cases} \liminf_{\rho\downarrow 0} \dfrac{\mu(B_\rho(x))}{\mu_0(B_\rho(x))} & \text{for } x \in X \setminus (U_0 \cup V_0) \\ \infty & \text{for } x \in U_0 \setminus V_0 \\ 0 & \text{for } x \in V_0, \end{cases}$$

with U_0, V_0 as in 3.9. Then there is an analogue of the Comparison Theorem 3.13 for the lower density. Preparatory to that we need the following lemma:

3.19 Lemma. *If μ, μ_0 is any pair of Borel regular measures on X with μ σ-finite, then there is a Borel set $B \subset X$ with $\mu_0(B) = 0$ and $\mu \llcorner (X \setminus B)$ absolutely continuous with respect to μ_0 (i.e. $\mu_0(S) = 0 \Rightarrow \mu(S \setminus B) = 0\ \forall S \subset X$).*

Proof: Case 1: $\mu(X) < \infty$. In this case let $\mathcal{A} = \{$Borel sets $A \subset X$ with $\mu_0(A) = 0\}$, $\alpha = \sup\{\mu(A) : A \in \mathcal{A}\}$, and choose a sequence $A_j \in \mathcal{A}$ with $\lim \mu(A_j) = \alpha$. Then $B = \cup_j A_j \in \mathcal{A}$ and if $S \subset X$ with $\mu_0(S) = 0$ then, by Borel regularity of μ_0, we can select $A \in \mathcal{A}$ with $S \subset A$ and $\mu(S \setminus B) \leq \mu(A \setminus B) = \mu(B \cup (A \setminus B)) - \mu(B) \leq \alpha - \alpha = 0$, so B has the required property.

Case 2: The general case when μ is σ-finite. In this case we can select an increasing sequence A_j of Borel sets with $X = \cup_j A_j$ and $\mu(A_j) < \infty$ for each j. Applying Case 1 to $\mu \llcorner A_j$, there is a Borel set B_j with $\mu_0(B_j) = 0$ and $\mu \llcorner (A_j \setminus B_j)$ absolutely continuous with respect to μ_0. So we can take $B = \cup_j B_j$. □

3.20 Remark: The set B is evidently unique up to a set of μ-measure zero, so the Borel regular measure $\mu \llcorner (X \setminus B)$ is uniquely determined; it is called the absolutely continuous part of μ relative to μ_0.

We can now prove an analogue of the Comparison Theorem 3.13 for the lower density:

3.21 Theorem. *Suppose μ, μ_0 are open σ-finite (as in* 1.14*) Borel regular measures on X, $t > 0$, and $A \subset X$ with $\Theta_*^{\mu_0}(\mu, x) \le t$ for all $x \in A$.*

(i) *If μ has the Symmetric Vitali Property then $\mu(A) \le t\mu_0(A)$.*

(ii) *If μ_0 has the Symmetric Vitali Property then $\mu(A \setminus B) \le t\mu_0(A)$, where B (with $\mu_0(B) = 0$) is as in* 3.19.

Proof: The proof is similar to the proof of Theorem 3.13. In view of the open σ-finiteness property we can suppose without loss of generality that both $\mu(X) < \infty$ and $\mu_0(X) < \infty$.

Proof of (i): First observe that $A \subset X \setminus U_0$ (because, by Definition 3.18, $\Theta_*^{\mu_0}(\mu, x) = \infty$ on U_0). Let $\tau > t$. By Theorem 1.15(1) we can select an open $U \supset A$ with $\mu_0(U) < \mu_0(A) + \tau - t$.

Define

$$\mathcal{B} = \{B_\rho(x) \subset U : x \in A \text{ and } \mu(B_\rho(x)) < \tau\mu_0(B_\rho(x))\}.$$

$\mathcal{B}$ evidently covers A finely, so by the Symmetric Vitali Property for μ there is a pairwise disjoint collection $B_{\rho_j}(x_j)$ with $\mu(A \setminus (\cup_j B_{\rho_j}(x_j))) = 0$ and $\mu(B_{\rho_j}(x_j)) \le \tau\mu_0(B_{\rho_j}(x_j))$ for each j. By summing on j we then have $\mu(A) \le \tau\mu_0(U) \le \tau(\mu_0(A) + \tau - t)$, so letting $\tau \downarrow 0$ gives the required result.

Proof of (ii): With B be as in Lemma 3.19, $\tilde{\mu} = \mu \llcorner (X \setminus B)$ is absolutely continuous with respect to μ_0, hence the Symmetric Vitali Property for μ_0 implies the Symmetric Vitali Property for $\tilde{\mu}$, so we can apply part (i) with $A \setminus B$ in place of A and $\tilde{\mu}$ in place of μ. This gives the required result. □

We define the density $\Theta^{\mu_0}(\mu, x)$ to be the common value of $\Theta^{*\mu_0}(\mu, x)$ and $\Theta_*^{\mu_0}(\mu, x)$ at points where these quantities are equal. Thus if U_0, V_0 are the open sets in 3.9 and 3.18, then

$$\textbf{3.22} \quad \Theta^{\mu_0}(\mu, x) = \begin{cases} \lim\limits_{\rho\downarrow 0} \dfrac{\mu(B_\rho(x))}{\mu_0(B_\rho(x))} & \text{if } x \in X \setminus (U_0 \cup V_0) \text{ and this limit exists} \\ \infty & \text{at points } x \in U_0 \setminus V_0 \\ 0 & \text{at points } x \in V_0, \end{cases}$$

and $\Theta^{\mu_0}(\mu, x)$ is undefined at points where $\Theta_*^{\mu_0}(\mu, x) < \Theta^{*\mu_0}(\mu, x)$.

3.23 Theorem (Differentiation Theorem.) *Suppose μ, μ_0 are open σ-finite (as in* 1.14*) Borel regular measures on X.*

(i) *If μ has the Symmetric Vitali Property, then there is a Borel set S of μ-measure zero such that $\Theta^{\mu_0}(\mu, x)$ (as in* 3.22*) exists for all $x \in X \setminus S$.*

(ii) *If μ_0 has the Symmetric Vitali Property, then there is a Borel set S of μ_0-measure zero such $\Theta^{\mu_0}(\mu, x)$ exists and is finite for all $x \in X \setminus S$.*

In either case $\Theta^{\mu_0}(\mu, x)$ is a Borel measurable function of $x \in X \setminus S$.

Proof: First assume $\mu_0(X), \mu(X) < \infty$ and let $A \subset X$ be any Borel set.

To prove (i) first note that by the Comparison Theorems 3.13 and 3.21(i), for any given $a, b > 0$,

(1) $$\Theta_*^{\mu_0}(\mu, x) < a \text{ and } \Theta^{*\mu_0}(\mu, x) > b \text{ for all } x \in A$$
$$\Rightarrow \mu(A) \le a\mu_0(A) \text{ and } b\mu_0(A) \le \mu(A).$$

In particular if $0 < a < b$ and

$$E_{a,b} = \{x \in X \setminus U_0 : \Theta_*^{\mu_0}(\mu, x) < a < b < \Theta^{*\mu_0}(\mu, x)\}.$$

then $a^{-1}\mu(E_{a,b}) \le \mu_0(E_{a,b}) \le b^{-1}\mu(E_{a,b})$, which implies that

(2) $$\mu_0(E_{a,b}) = \mu(E_{a,b}) = 0.$$

Since $\{x : \Theta_*^{\mu_0}(\mu, x) < \Theta^{*\mu_0}(\mu, x)\} = \cup_{a,b \text{ rational}, a<b} E_{a,b}$ we deduce from (2) that $\Theta_*^{\mu_0}(\mu, x) = \Theta^{*\mu_0}(\mu, x)$ for μ_0-a.e. $x \in X \setminus U_0$, so indeed $\Theta^{\mu_0}(\mu, x)$ exists and is in $[0, \infty]$ for μ-a.e. $x \in X \setminus U_0$. $\Theta^{\mu_0}(\mu, x)$ is also defined in U_0 by Definition 3.22. Thus $\Theta^{\mu_0}(\mu, x)$ is well-defined μ-a.e., so by Borel regularity of μ there is a Borel set S with $\mu(S) = 0$ such that $\Theta^{\mu_0}(\mu, x)$ is well-defined for all $x \in X \setminus S$.

The measurability of $\Theta^{\mu_0}(\mu, x)$ as a function of $x \in X \setminus S$ is proved as follows: For each fixed $\rho > 0$, $\mu(B_\rho(x))$ and $\mu_0(B_\rho(x))$ are positive upper semi-continuous functions of $x \in X \setminus (S \cup V_0 \cup U_0)$, hence are Borel measurable functions on $X \setminus (S \cup V_0 \cup U_0)$, and hence so is the quotient $\mu(B_\rho(x))/\mu_0(B_\rho(x))$. Hence $\Theta^{\mu_0}(\mu, x) = \lim_{i\to\infty} \mu(B_{1/i}(x))/\mu_0(B_{1/i}(x))$ is Borel measurable on $X \setminus (S \cup V_0 \cup U_0)$. Finally, by Definition 3.22, $\Theta^{\mu_0}(\mu, x) = \infty$ on $U_0 \setminus V_0$ and $\Theta^{\mu_0}(\mu, x) = 0$

on V_0. Since U_0, V_0 are open we then conclude that indeed $\Theta^{\mu_0}(\mu, x)$ is Borel measurable in case μ, μ_0 are finite measures. In the general open σ-finite case, when there are open sets V_j with $\cup_j V_j = X$ and $\mu(V_j), \mu_0(V_j) < \infty$, we apply the above with $\mu \llcorner V_j, \mu_0 \llcorner V_j$ in place of μ, μ_0 respectively.

To prove (ii), note first that by Corollary 3.14 we have

$$(3) \qquad \Theta^{*\mu_0}(\mu, x) < \infty \text{ for } \mu_0\text{-a.e. } x \in X.$$

As in 3.19, let B be a Borel set of μ_0-measure zero such that $\widetilde{\mu} = \mu \llcorner (X \setminus B)$ is absolutely continuous with respect to μ_0. Then $\widetilde{\mu}$ has the Symmetric Vitali Property, and hence the argument of (i) above applies with $\widetilde{\mu}$ in place of μ to give

$$(4) \qquad \mu_0(E_{a,b}) = \mu(E_{a,b} \setminus B) = 0,$$

in place of (2). Hence $\Theta^{\mu_0}(\mu, x)$ exists for μ_0-a.e. $x \in X$, and by (3) it is also finite for μ_0-a.e. $x \in X$, hence there is a Borel set S with $\mu_0(S) = 0$ such that $\Theta^{\mu_0}(\mu, x)$ exists and is finite for all $x \in X \setminus S$.

The measurability of $\Theta^{\mu_0}(\mu, x)$ follows similarly to case (i) above. □

Next, recall the abstract Radon-Nikodym theorem, which says that if μ, μ_0 are abstract σ-finite measures on a σ-algebra $\mathcal{A}$ of subsets of an abstract space X, and if μ is absolutely continuous with respect to μ_0 (i.e. $A \in \mathcal{A}$ with $\mu_0(A) = 0 \Rightarrow \mu(A) = 0$), then there is a non-negative $\mathcal{A}$–measurable function θ on X such that

$$\mu(A) = \int_A \theta \, d\mu_0, \quad A \in \mathcal{A}.$$

In these circumstances the function θ is called "the Radon-Nikodym derivative" of μ with respect to μ_0, denoted $\frac{d\mu}{d\mu_0}$ or $D_{\mu_0}\mu$.

We show here that in case μ, μ_0 are Borel regular open σ-finite (as in 1.14) on the metric space X with μ_0 having the Symmetric Vitali Property, then the Radon-Nikodym derivative $D_{\mu_0}\mu(x)$ is just the density $\Theta^{\mu_0}(\mu, x) = \lim_{\rho \downarrow 0} \frac{\mu(B_\rho(x))}{\mu_0(B_\rho(x))}$:

3.24 Theorem (Radon-Nikodym.) *Suppose μ, μ_0 are open σ-finite (as in* 1.14*) Borel regular measures on X, and μ_0 has the Symmetric Vitali Property.*

(i) *If μ is absolutely continuous with respect to μ_0 (i.e. $E \subset X$ with $\mu_0(E) = 0 \Rightarrow \mu(E) = 0$ and hence μ also has the Symmetric Vitali Property), then*

$$(*) \qquad \mu(A) = \int_A \Theta^{\mu_0}(\mu, x) \, d\mu_0(x) \textit{ for every Borel set } A \subset X.$$

(ii) *If we drop the condition that μ is absolutely continuous with respect to μ_0, then in place of $(*)$ we can still conclude that there is a Borel set Z with $\mu_0(Z) = 0$ and*

$$(\ddagger) \qquad \mu(A) = \int_A \Theta^{\mu_0}(\mu, x) \, d\mu_0(x) + (\mu \llcorner Z)(A),$$

for each Borel set $A \subset X$.

(iii) *Finally, if μ also has the Symmetric Vitali Property, then we get $(\ddagger)$ with*

$$Z = \{x \in X : \Theta^{\mu_0}(\mu, x) = \infty\}$$

(which is a set of μ_0-measure zero by 3.23(ii)*).*

3.25 Remarks: (1) By Remark 3.11(3) we always have the conclusion of 3.23(iii) if $X = \mathbb{R}^n$.

(2) $\mu \llcorner Z$ is called the singular part of μ with respect to μ_0.

Proof of Theorem 3.24: Since μ, μ_0 are open σ-finite, we can assume $\mu(X) < \infty, \mu_0(X) < \infty$. Let S be a Borel set of μ_0-measure zero as in Theorem 3.23. For any Borel set $A \subset X \setminus S$ let

$$\nu(A) = \int_A \Theta^{\mu_0}(\mu, x) \, d\mu_0(x)$$

and for any subset $A \subset X \setminus S$ let $\nu(A) = \inf_{B \supset A, B \text{ Borel}} \nu(B)$. By Remark 1.11, ν is a Radon measure and, with $0 < a < b$, $A_{a,b} = \{x \in A : a < \Theta^{\mu_0}(\mu, x) < b\}$ and A any Borel set, we have

$$a\mu_0(A_{a,b}) \le \nu(A_{a,b}) \le b\mu_0(A_{a,b}).$$

On the other hand the Comparison Theorems 3.13, 3.21(i) imply

$$a\mu_0(A_{a,b}) \le \mu(A_{a,b}) \le b\mu_0(A_{a,b}),$$

and so

$$\frac{a}{b}\mu_0(A_{a,b}) \le \nu(A_{a,b}) \le \frac{b}{a}\mu_0(A_{a,b})$$

and it follows that $\nu(A) = \mu(A)$. Thus $(*)$ is proved.

In the general case (when we allow the possibility that there are sets A with $\mu_0(A) = 0$ and $\mu(A) > 0$), we can apply the previous argument to the Borel regular measure $\widetilde{\mu} = \mu \llcorner (X \setminus B)$, where B is the set of μ_0-measure zero of Lemma 3.19. This gives

$$\mu(A \setminus B) = \int_A \Theta^{\mu_0}(\mu, x) \, d\mu_0 \ \forall \text{ Borel set } A \subset X.$$

Thus 3.23$(\ddagger)$ holds with $Z = B$.

Finally, in case μ also has the Symmetric Vitali Property, Theorem 3.23(i) establishes that $\Theta^{\mu_0}(\mu, x)$ exists μ-almost everywhere (as well as μ_0-almost everywhere) in X. On the other hand if $\widetilde{X} = X \setminus U_0$ and $A \subset \{x \in \widetilde{X} : \Theta^{\mu_0}(\mu, x) < \infty\}$ ($= \cup_{n=1}^\infty \{x \in \widetilde{X} : \Theta^{\mu_0}(\mu, x) < n\}$) then by Theorem 3.21(i)

$$\mu_0(A) = 0 \Rightarrow \mu(A) = 0,$$

and we can therefore apply $(*)$ with $\mu \llcorner (X \setminus Z)$, $Z = \{x : \Theta^{\mu_0}(\mu, x) = \infty\}$, in place of μ. Hence (iii) is proved. □

We conclude this section with two important bounds for densities with respect to Hausdorff measure.

3.26 Theorem. *For any $\mathcal{H}^n$-measurable subset of A of X:*

(1) *If $\mathcal{H}^n(A) < \infty$, then $\Theta^{*n}(\mathcal{H}^n, A, x) \le 1$ for $\mathcal{H}^n$-a.e. $x \in A$.*

(2) *If $\mathcal{H}^n_\delta(A) < \infty$ for each $\delta > 0$ (note this is automatic if A is a totally bounded subset of X), then $\Theta^{*n}(\mathcal{H}^n_\infty, A, x) \ge 2^{-n}$ for $\mathcal{H}^n$-a.e. $x \in A$.*

3.27 Remark: Since $\mathcal{H}^n \ge \mathcal{H}^n_\delta \ge \mathcal{H}^n_\infty$ (by Definitions 2.1, 2.2) this theorem implies

$$2^{-n} \le \Theta^{*n}(\mathcal{H}^n, A, x) \le 1 \text{ for } \mathcal{H}^n\text{-a.e. } x \in A,$$

provided $\mathcal{H}^n(A) < \infty$.

Proof of 3.26: To prove (1), let $\varepsilon, t > 0$, let $A_t = \{x \in A : \Theta^{*n}(\mathcal{H}^n, A, x) \ge t\}$ and (using 1.15(1) with $\mu = \mathcal{H}^n \llcorner A$), choose an open set $U \supset A_t$ such that

$$\mathcal{H}^n(U \cap A) < \mathcal{H}^n(A_t) + \varepsilon.$$

Since U is open and since $A_t \subset U$ we have $\Theta^{*n}(\mathcal{H}^n, A \cap U, x) \ge t$ for each $x \in A_t$. Hence 3.3 (with $\mathcal{H}^n \llcorner A, A_t, A \cap U$ in place of μ, A_1, A_2) implies that

$$t\mathcal{H}^n(A_t) \le \mathcal{H}^n(A \cap U) \le \mathcal{H}^n(A_t) + \varepsilon.$$

We thus have $\mathcal{H}^n(A_t) = 0$ for each $t > 1$. Since $\{x : \Theta^{*n}(\mathcal{H}^n, A, x) > 1\} = \cup_{j=1}^\infty A_{t_j}$ for any decreasing sequence $\{t_j\}$ with $\lim t_j = 1$, we thus have $\mathcal{H}^n\{x : \Theta^{*n}(\mathcal{H}^n, A, x) > 1\} = 0$, as required.

To prove (2), suppose for contradiction that $\Theta^{*n}(\mathcal{H}^n_\infty \llcorner A, x) < 2^{-n}$ for each x in a set $B_0 \subset A$ with $\mathcal{H}^n(B_0) > 0$. Then for each $x \in B_0$ select $\delta_x \in (0,1)$ such that

$$\mathcal{H}^n_\infty(A \cap B_\rho(x)) \le \frac{1-\delta_x}{2^n}\omega_n \rho^n, \ 0 < \rho < \delta_x.$$

Therefore, since $B_0 = \cup_{j=1}^\infty \{x \in B_0 : \delta_x > 1/j\}$ and since $\mathcal{H}^n_\delta(A \cap B_\rho(x)) \equiv \mathcal{H}^n_\infty(A \cap B_\rho(x))$ for any $\rho < \delta/2$ (by Definition 2.2), we can select $\delta > 0$ and $B \subset B_0$ with $\mathcal{H}^n(B) > 0$ and

$$(1) \qquad \mathcal{H}^n_\delta(A \cap B_\rho(x)) \le \frac{1-\delta}{2^n}\omega_n \rho^n, \ 0 < \rho < \delta/2, \ x \in B.$$

XXX Now using 2.2 again, we can choose sets $C_1, C_2, \ldots$ with $B \subset \cup_{j=1}^\infty C_j$, $C_j \cap B \neq \emptyset$, $\operatorname{diam} C_j < \delta \ \forall j$, and

$$(2) \qquad \textstyle\sum_j \omega_n \rho_j^n < \frac{1}{1-\delta}\mathcal{H}^n_\delta(B), \ \rho_j = \operatorname{diam} C_j/2$$

Now take $x_j \in C_j \cap B$, so that $B \subset A \cap (\cup_{j=1}^\infty B_{2\rho_j}(x_j))$, and we conclude from (1), (2) that $\mathcal{H}^n_\delta(B) = 0$, hence $\mathcal{H}^n(B) = 0$, contradicting our choice of B. □

4 Radon Measures, Representation Theorem

In this section we work mainly in locally compact Hausdorff spaces, and for the reader's convenience we recall some basic definitions and preliminary topological results for such spaces.

Recall that a topological space is said to be Hausdorff if it has the property that for every pair of distinct points $x, y \in X$ there are open sets U, V with $x \in U$, $y \in V$ and $U \cap V = \emptyset$. In such a space *all compact sets are automatically closed*, the proof of which is as follows: observe that if $x \notin K$ then for each $y \in K$ we can (by definition of Hausdorff space) pick open U_y, V_y with $x \in U_y$, $y \in V_y$ and $U_y \cap V_y = \emptyset$. By compactness of K there is a finite set $y_1, \ldots, y_N \in K$ with $K \subset \cup_{j=1}^N V_{y_j}$. But then $\cap_{j=1}^N U_{y_j}$ is an open set containing x which is disjoint from $\cup_j V_{y_j}$ and hence disjoint from K, so that K is closed as claimed. In fact we proved a bit more: that for each $x \notin K$ there are disjoint open sets U, V with $x \in U$ and $K \subset V$. Then if L is another compact set disjoint from K we can repeat this for each $x \in L$ thus obtaining disjoint open U_x, V_x with $x \in U_x$ and $K \subset V_x$, and then compactness of L implies $\exists x_1, \ldots, x_M \in L$ such that $L \subset \cup_{j=1}^M U_{x_j}$ and then $\cup_{j=1}^M U_{x_j}$ and $\cap_{j=1}^M V_{x_j}$ are disjoint open sets containing L and K respectively. By a simple inductive argument (left as an exercise) we can extend this to finite pairwise disjoint unions of compact subsets:

4.1 Lemma. *Let X be a Hausdorff space and $K_1, \ldots, K_N$ be pairwise disjoint compact subsets of X. Then there are pairwise disjoint open subsets $U_1, \ldots, U_N$ with $K_j \subset U_j$ for each $j = 1, \ldots, N$.*

Notice in particular that we have the following corollary of Lemma 4.1:

4.2 Corollary. *A compact Hausdorff space is normal: i.e. given closed disjoint subsets K_1, K_2 of a compact Hausdorff space, we can find disjoint open U_1, U_2 with $K_j \subset U_j$ for $j = 1, 2$.*

Most of the rest of the discussion here takes place in locally compact Hausdorff space: A space X is said to be *locally compact* if for each $x \in X$ there is a neighborhood U_x of x such that the closure $\overline{U}_x$ of U_x is compact.

An important preliminary lemma in such spaces is:

4.3 Lemma. *If X is a locally compact Hausdorff space and V is a neighborhood of a* *point x, then there is a neighborhood U_x of x such that $\overline{U}_x$ is a compact subset of V.*

Proof: First pick a neighborhood W_0 of x such that $\overline{W}_0$ is compact and define $W = W_0 \cap V$. Then $\overline{W}$ is compact and hence, with the subspace topology, is normal by Corollary 1 above. Hence since $\overline{W} \setminus W$ and $\{x\}$ are disjoint closed sets in this space, and since open sets in the subspace $\overline{W}$ can by definition be expressed as the intersection of open sets from X with the subset $\overline{W}$, we can find open U_1, U_2 in the space X with $x \in U_1$, $\overline{W} \setminus W \subset U_2$ and $U_1 \cap U_2 \cap \overline{W} = \emptyset$. The last identity says $U_1 \cap \overline{W} \subset \overline{W} \setminus U_2$, whence $x \in U_1 \cap W \subset \overline{W} \setminus U_2 \subset W \subset V$, and since $\overline{W} \setminus U_2$ is a closed set, we then have $x \in U_1 \cap W \subset \overline{U_1 \cap W} \subset \overline{W} \setminus U_2 \subset V$, so the lemma is proved with $U_x = U_1 \cap W$. □

Remark: In locally compact Hausdorff space, using Lemmas 4.1 and 4.3 it is easy to check that we can select the U_j in Lemma 4.1 above to have compact pairwise disjoint closures.

The following lemma is a version of the Urysohn lemma valid in locally compact Hausdorff space:

4.4 Lemma. *Let X be a locally compact Hausdorff space, $K \subset X$ compact, and $K \subset V$, V open. Then there is an open $U \supset K$ with $\overline{U} \subset V$, $\overline{U}$ compact, and an $f : X \to [0,1]$ with $f \equiv 1$ in a neighborhood of K and $f \equiv 0$ on $X \setminus U$.*

Proof: By Lemma 4.3 each $x \in K$ has a neighborhood U_x with $\overline{U}_x \subset V$. Then by compactness of K we have $K \subset U \equiv \cup_{j=1}^N U_{x_j}$ for some finite collection $x_1, \ldots, x_N \in K$ and $\overline{U} = \cup_{j=1}^N \overline{U}_{x_j} \subset V$. Now $\overline{U}$ is compact, so by Corollary 1 it is a normal space and the Urysohn lemma can be applied to give $f_0 : \overline{U} \to [0,1]$ with $f_0 \equiv 1$ on K and and $f_0 \equiv 0$ on $\overline{U} \setminus U$. Then of course the function f_1 defined by $f_1 \equiv f_0$ on $\overline{U}$ and $f_1 \equiv 0$ on $X \setminus \overline{U}$ is continuous (check!) because $f|\overline{U}$ is continuous and f is identically zero (the value of $f|X \setminus \overline{U}$) on the overlap set $\overline{U} \setminus U \equiv \overline{U} \cap (X \setminus U)$. Finally we let $f \equiv 2\min\{f_1, \frac{1}{2}\}$ and observe that f is then identically 1 in the set where $f_1 > \frac{1}{2}$, which is an open set containing K, and f evidently has all the remaining stated properties. □

The following corollary of Lemma 4.4 is important:

4.5 Corollary (Partition of Unity.) *If X is a locally compact Hausdorff space, $K \subset X$ is compact, and if $U_1, \ldots, U_N$ is any open cover for K, then there exist continuous $\varphi_j : X \to [0,1]$ such that support φ_j is a compact subset of U_j for each j, and $\sum_{j=1}^N \varphi_j \equiv 1$ in a neighborhood of K.*

Proof: By Lemma 4.3, for each $x \in K$ there is a $j \in \{1, \ldots, N\}$ and a neighborhood U_x of x such that $\overline{U}_x$ is a compact subset of this U_j. By compactness of K we have finitely many of these neighborhoods, say $U_{x_1}, \ldots, U_{x_N}$ with $K \subset \cup_{i=1}^N U_{x_i}$. Then for each $j = 1, \ldots, N$ we define V_j to be the union of all U_{x_i} such that $\overline{U}_{x_i} \subset U_j$. Then the $\overline{V}_j$ is a compact subset of U_j for each j, and the V_j cover K. So by Lemma 4.4 for each $j = 1, \ldots, N$ we can select $\psi_j : X \to [0,1]$ with $\psi_j \equiv 1$ on $\overline{V}_j$ and $\psi_j \equiv 0$ on $X \setminus W_j$ for some open W_j with $\overline{W}_j$ a compact subset of U_j and $W_j \supset \overline{V}_j$. We can also use Lemma 4.4 to select $f_0 : X \to [0,1]$ with $f_0 \equiv 1$ in the neighborhood $\cup_{j=1}^N V_j$ of K and $f_0 \equiv 0$ on $\{x : \sum_{j=1}^N \psi_j(x) = 0\}$. (This latter set is closed and has (open) complement which is a neighborhood of the compact set $\cup_{j=1}^N \overline{V}_j$ and so we can indeed construct such f_0 by Lemma 4.4.) Then set $\psi_0 = 1 - f_0$ and observe that by construction $\sum_{i=0}^N \psi_i > 0$ everywhere on X, so we can define continuous functions φ_j by

$$\varphi_j = \frac{\psi_j}{\sum_{i=0}^N \psi_i}, \quad j = 1, \ldots, N.$$

Evidently these functions have the required properties. □

We now give the definition of Radon measure. Radon measures are typically used only in locally compact Hausdorff space, but the definition and the first two lemmas following it are valid in arbitrary Hausdorff space:

4.6 Definition: Given a Hausdorff space X, a "Radon measure" on X is an outer measure μ on X having the 3 properties:

$$\mu \text{ is Borel regular and } \mu(K) < \infty \;\; \forall \text{ compact } K \subset X \tag{R1}$$

$$\mu(A) = \inf_{U \text{ open}, U \supset A} \mu(U) \text{ for each subset } A \subset X \tag{R2}$$

$$\mu(U) = \sup_{K \text{ compact}, K \subset U} \mu(K) \text{ for each open } U \subset X. \tag{R3}$$

Such measures automatically have a property like (R3) with an arbitrary μ-measurable subset of finite measure:

4.7 Lemma. *Let X be a Hausdorff space and μ a Radon measure on X. Then μ automatically has the property*

$$\mu(A) = \sup_{K \subset A, K \text{ compact}} \mu(K)$$

for every μ-measurable set $A \subset X$ with $\mu(A) < \infty$.

Proof: Let $\varepsilon > 0$. By definition of Radon measure we can choose an open U containing A with $\mu(U \setminus A) < \varepsilon$, and then a compact $K \subset U$ with $\mu(U \setminus K) < \varepsilon$

and finally an open W containing $U \setminus A$ with $\mu(W \setminus (U \setminus A)) < \varepsilon$ (so that $\mu(W) \le \varepsilon + \mu(U \setminus A) < 2\varepsilon$). Then we have that $K \setminus W$ is a compact subset of $U \setminus W$, which is a subset of A, and also

$$\mu(A \setminus (K \setminus W)) \le \mu(U \setminus (K \setminus W)) \le \mu(U \setminus K) + \mu(W) \le 3\varepsilon,$$

which completes the proof. □

The following lemma asserts that the defining property (R1) of Radon measures follows automatically from the remaining two properties ((R2) and (R3)) in case μ is finite and additive on finite disjoint unions of compact sets.

4.8 Lemma. *Let X be a Hausdorff space and assume that μ is an outer measure on X satisfying the properties* (R2), (R3) *above, and in addition assume that*

$$\mu(K_1 \cup K_2) = \mu(K_1) + \mu(K_2) < \infty \text{ whenever } K_1, K_2 \text{ are compact and disjoint.}$$

Then (R1) *holds and hence μ is a Radon measure.*

Proof: Note that (R2) implies that for every set $A \subset X$ we can find open sets U_j such that $A \subset \cap_j U_j$ and $\mu(A) = \mu(\cap_j U_j)$. So to complete the proof of (R1) we just have to check that all Borel sets are μ-measurable; since the μ-measurable sets form a σ-algebra and the Borel sets form the smallest σ-algebra of subsets of X which contains all the open sets, we thus need only to check that all open sets are μ-measurable.

Let $\varepsilon > 0$ be arbitrary, Y an arbitrary subset of X with $\mu(Y) < \infty$ and let U be an arbitrary open subset of X. By (R2) we can pick an open set $V \supset Y$ with $\mu(V) < \mu(Y) + \varepsilon$ and by (R3) we can pick a compact set $K_1 \subset V \cap U$ with $\mu(V \cap U) \le \mu(K_1) + \varepsilon$, and then a compact set $K_2 \subset V \setminus K_1$ with $\mu(V \setminus K_1) \le \mu(K_2) + \varepsilon$. Then

$$\begin{aligned}\mu(V \setminus U) + \mu(V \cap U) &\le \mu(V \setminus K_1) + \mu(K_1) + \varepsilon \\ &\le \mu(K_2) + \mu(K_1) + 2\varepsilon \\ &= \mu(K_2 \cup K_1) + 2\varepsilon \text{ (by (i))} \\ &\le \mu((V \setminus K_1) \cup K_1) + 2\varepsilon = \mu(V) + 2\varepsilon \le \mu(Y) + 3\varepsilon,\end{aligned}$$

hence $\mu(Y \setminus U) + \mu(Y \cap U) \le \mu(V \setminus U) + \mu(V \cap U) \le \mu(Y) + 3\varepsilon$ which by arbitrariness of ε gives $\mu(Y \setminus U) + \mu(Y \cap U) \le \mu(Y)$, which establishes the μ-measurability of U. Thus all open sets are μ-measurable, and hence all Borel sets are μ-measurable, and so (R1) is established. □

The following lemma guarantees the convenient fact that, in a locally compact space

such that all open subsets are σ-compact, all locally finite Borel regular outer measures are in fact Radon measures.

4.9 Lemma. *Let X be a locally compact Hausdorff space and suppose that each open set is the countable union of compact subsets. Then any Borel regular outer measure on X which is finite on each compact set is automatically a Radon measure.*

Proof: First observe that in a Hausdorff space X the statement "each open set is the countable union of compact subsets" is equivalent to the statement "X is σ-compact (i.e. the countable union of compact sets) and every closed set is the countable intersection of open sets" as one readily checks by using De Morgan's laws and the fact that a set is open if and only if its complement is closed. Thus we have at our disposal the facts that X is σ-compact and every closed set is a countable intersection of open sets. The latter fact enables us to apply the Theorem 1.15 on Borel regular outer measures, and we can therefore assert that

(1) $\quad \mu(A) = \inf_{U \text{ open}, A \subset U} \mu(U)$ whenever $A \subset X$ has the property
$\exists$ open V_j with $A \subset \cup_j V_j$ and $\mu(V_j) < \infty \ \forall j$

and

(2) $\quad \mu(A) = \sup_{C \text{ closed}, C \subset A} \mu(C)$, provided $A = \cup_j A_j$ with
A_j is μ-measurable and $\mu(A_j) < \infty \ \forall j$.

Now observe that in a locally compact Hausdorff space it is true that for each compact $K \subset X$ we can always find an open set $V \supset K$ such that $\overline{V}$ (the closure of V) is compact. (This easily follows from the definition of compactness and the fact that each point of K has an open neighborhood with compact closure.) If $X = \cup_{j=1}^{\infty} K_j$, where each K_j is compact, that we can apply this with K_j in place of K, and we deduce that there are open sets V_j in X such that $\cup_j V_j = X$ and $\mu(V_j) < \infty$ for each j, and so in this case (when X is σ-compact) the identity in (1) holds for every subset $A \subset X$; that is

$$\mu(A) = \inf_{U \text{ open}, A \subset U} \mu(U) \text{ for every } A \subset X,$$

which is the property (R2). Next we note that if $A \subset X$ is μ-measurable, then we can write $A = \cup_j A_j$, where $A_j = A \cap K_j$ (because $X = \cup_j K_j$) and $\mu(A_j) \le \mu(K_j) < \infty$ for each j, so (2) actually holds for every μ-measurable A in case X is σ-compact (i.e. in case $X = \cup_{j=1}^{\infty} K_j$ with K_j compact), and for any closed set C we can write $C = \cup_j C_j$ where C_j is the increasing sequence of compact sets given by $C_j = C \cap (\cup_{i=1}^{j} K_i)$ and so $\mu(C) = \lim_j \mu(C_j)$ and hence $\mu(C) =$

$\sup_{K\subset C,\, K \text{ compact}} \mu(K)$. Thus in the σ-compact case (2) actually tells us that $\mu(A) = \sup_{K\subset A,\, K \text{ compact}} \mu(K)$ for any μ-measurable set A. This in particular holds for $A =$ an open set, which is the remaining property (R3) we needed. □

Next we have the following important density result:

4.10 Theorem. *Let X be a locally compact Hausdorff space, μ a Radon measure on X and $1 \le p < \infty$. Then $C_c(X)$ is dense in $L^p(\mu)$; that is, for each $\varepsilon > 0$ and each $f \in L^p$ there is a $g \in C_c(X)$ such that $\|g - f\|_p < \varepsilon$.*

In view of Remark 1.11 and Lemma 4.9 we see that Theorem 4.10 directly implies the following:

4.11 Corollary. *If X is a locally compact Hausdorff space such that every open set in X is the countable union of compact sets, and if μ is any Borel measure on X which is finite on each compact set, then the space $C_c(X)$ is dense in $L^1(\mu)$ and μ is the restriction to the Borel sets of a Radon measure $\overline{\mu}$.*

Proof of Theorem 4.10: Let $f : X \to \mathbb{R}$ be μ-measurable with $\|f\|_p < \infty$ and let $\varepsilon > 0$. Observe that the simple functions are dense in $L^p(\mu)$ (which one can check using the dominated convergence theorem and the fact that both f_+ and f_- can be expressed as the pointwise limits of increasing sequences of non-negative simple functions), so we can pick a simple function $\varphi = \sum_{j=1}^N a_j \chi_{A_j}$, where the a_j are distinct non-zero reals and A_j are pairwise disjoint μ-measurable subsets of X, such that $\|f - \varphi\|_p < \varepsilon$. Since $\|\varphi\|_p \le \|\varphi - f\|_p + \|f\|_p < \infty$ we must then have $\mu(A_j) < \infty$ for each j. Pick $M > \max\{|a_1|, \ldots, |a_N|\}$ and use Lemma 4.7 to select compact $K_j \subset A_j$ with $\mu(A_j \setminus K_j) < \varepsilon^p/(2^{p+1}M^pN)$. Also, using the definition of Radon measure, we can find open $U_j \supset K_j$ with $\mu(U_j \setminus K_j) < \varepsilon^p/(2^{p+1}M^pN)$ and by Lemma 4.7 we can assume without loss of generality that these open sets $U_1, \ldots, U_N$ are pairwise disjoint (otherwise replace U_j by $U_j \cap U_j^0$, where $U_1^0, \ldots, U_N^0$ are pairwise disjoint open sets with $K_j \subset U_j^0$). By Lemma 4.4 we have $g_j \in C_c(X)$ with $g_j \equiv a_j$ on K_j, $\{x : g_j(x) \neq 0\}$ contained in a compact subset of U_j, and $\sup|g_j| \le |a_j|$, and hence by the pairwise disjointness of the U_j we have that $g \equiv \sum_{j=1}^N g_j$ agrees with φ on each K_j and $\sup|g| = \sup|\varphi| < M$. Then $\varphi - g$ vanishes off the set $\cup_j((U_j \setminus K_j) \cup (A_j \setminus K_j))$ and we have $\int_X |\varphi - g|^p \, d\mu \le \sum_j \int_{(U_j\setminus K_j)\cup(A_j - K_j)} |\varphi - g|^p \, d\mu \le (2M)^p \sum_j (\mu(A_j \setminus K_j) + \mu(U_j \setminus K_j)) \le \varepsilon^p$, and hence $\|f - g\|_p \le \|f - \varphi\|_p + \|\varphi - g\|_p \le 2\varepsilon$, as required.

We now state the Riesz representation theorem for non-negative functionals on the space $\mathcal{K}_+$, where, here and subsequently, $\mathcal{K}_+$ denotes the set of non-negative $C_c(X, \mathbb{R})$ functions, i.e. the set of continuous functions $f : X \to [0, \infty)$ with compact support.

4.12 Theorem (Riesz for non-negative functionals.) *Suppose X is a locally compact Hausdorff space, $\lambda : \mathcal{K}_+ \to [0, \infty)$ with $\lambda(cf) = c\lambda(f)$, $\lambda(f + g) = \lambda(f) + \lambda(g)$ whenever $c \ge 0$ and $f, g \in \mathcal{K}_+$, where $\mathcal{K}_+$ is the set of all non-negative continuous functions f on X with compact support. Then there is a Radon measure μ on X such that $\lambda(f) = \int_X f \, d\mu$ for all $f \in \mathcal{K}_+$.*

Before we begin the proof of 4.12 we observe the following 2 facts about the functional λ:

4.13 Remarks (1): Observe that if $f, g \in \mathcal{K}_+$ with $f \le g$ then $g - f \in \mathcal{K}_+$ and hence $\lambda(g) = \lambda(f + (g - f)) = \lambda(f) + \lambda(g - f) \ge \lambda(f)$.

(2) If K is compact, support $f \subset K$ and if $g \in \mathcal{K}_+$ with $g \equiv 1$ on K then we have $f \le (\sup f)g$ and $f\,g \equiv f$, so by Remark (1) above we have

$$(*) \qquad \lambda(f) \le (\sup f)\,\lambda(g), \quad f \in \mathcal{K}_+, \text{ support } f \subset K.$$

Notice in particular that if U is an arbitrary neighborhood of K then we can by Lemma 4.4 select neighborhood W of K with $\overline{W}$ a compact subset of U and a $g \in \mathcal{K}_+$ with $g \equiv 1$ in a neighborhood of $\overline{W}$, $g \le 1$ everywhere, and support $g \subset U$, whence the above inequality with $\overline{W}$ in place of K implies

$$(**) \qquad \sup_{f\in\mathcal{K}_+, f\le 1, \text{ support } f\subset W} \lambda(f) \le \inf_{g\in\mathcal{K}_+, g\le 1, g\equiv 1 \text{ in a nhd. of } \overline{W}, \text{ support } g\subset U} \lambda(g).$$

Proof of Theorem 4.12: For $U \subset X$ open, we define

$$(1) \qquad \mu(U) = \sup_{f\in\mathcal{K}_+, f\le 1, \text{support} f\subset U} \lambda(f),$$

and for arbitrary $A \subset X$ we define

$$(2) \qquad \mu(A) = \inf_{U\supset A, U \text{ open}} \mu(U).$$

Notice that these definitions are consistent when A is itself open. Notice also that by $(**)$ we have $\mu(K) < \infty$ for each compact K; indeed $(**)$ and the definitions (1), (2) evidently imply

$$(3) \qquad \mu(K) = \inf_{g\in\mathcal{K}_+, g\le 1, g\equiv 1 \text{ in a nhd. of } K} \lambda(g) \text{ for each compact } K \subset X,$$

Next we prove that μ is an outer measure. To see this, first let $U_1, U_2, \ldots$ be open and $U = \cup_j U_j$, then for any $f \in \mathcal{K}_+$ with $\sup f \le 1$ and support $f \subset U$ we have, by compactness of support f, that support $f \subset \cup_{j=1}^N U_j$ for some integer N, and by using a partition of unity $\varphi_1, \ldots, \varphi_N$ for support f subordinate to $U_1, \ldots, U_N$ (see the Corollary to Lemma 4.4 above), we have $\lambda(f) = \sum_{j=1}^N \lambda(\varphi_j f) \le \sum_{j=1}^N \mu(U_j)$. Taking sup over all such f we then have $\mu(U) \le \sum_j \mu(U_j)$. It then easily follows that $\mu(\cup_j A_j) \le \sum_j \mu(A_j)$ for each j. Since we trivially also have $\mu(\emptyset) = 0$ and $A \subset B \Rightarrow \mu(A) \le \mu(B)$ we thus have that μ is an outer measure on X.

Finally we want to show that μ is a Radon measure. For this we are going to use Lemma 4.8, so we have to check (R2), (R3) and the additivity property $\mu(K_1 \cup K_2) = \mu(K_1) + \mu(K_2)$ whenever K_1, K_2 are disjoint compact sets. But hypothesis (R2), (R3) are true by the definitions (1), (2), so we only have to check the the additivity on disjoint compact sets. In fact if K_1 and K_2 are disjoint compact subsets then for $\varepsilon > 0$ we can use (3) to find $g \in \mathcal{K}_+$ with $g \le 1$, $g \equiv 1$ in a neighborhood W of $K_1 \cup K_2$, and with $\lambda(g) \le \mu(K_1 \cup K_2) + \varepsilon$. By Lemma 4.1 we can then select disjoint open U_1, U_2 with $K_1 \subset U_1$ and $K_2 \subset U_2$, and by Lemma 4.4 we can select $f_1, f_2 \in \mathcal{K}_+$ with $f_j \equiv 1$ in a neighborhood of K_j such that support f_j is a compact subset of U_j and $f_j \le 1$ everywhere, $j = 1, 2$. Then by (3) $\mu(K_1) + \mu(K_2) \le \lambda(f_1 \cdot g) + \lambda(f_2 \cdot g) = \lambda((f_1 + f_2) \cdot g) \le \lambda(g) \le \mu(K_1 \cup K_2) + \varepsilon$. Thus $\mu(K_1) + \mu(K_2) \le \mu(K_1 \cup K_2)$, and of course the reverse inequality holds by subadditivity of μ, hence the hypotheses of Lemma 4.8 are all established and μ is a Radon measure.

Next observe that by $(*)$ we have $\lambda(h) \le \mu(\text{support}\, h) \sup h$, $h \in \mathcal{K}_+$, and hence (observing that h is the uniform limit of $\max\{h - 1/n, 0\}$ in X) we have

$$\lambda(h) \le \mu(\{x : h(x) > 0\}) \sup h, \ h \in \mathcal{K}_+. \tag{4}$$

For $f \in \mathcal{K}_+$ and $\varepsilon > 0$, we can select points $0 = t_0 < t_1 < t_2 < \ldots < t_{N-1} < \sup f < t_N$ with $t_j - t_{j-1} < \varepsilon$ for each $j = 1, \ldots, N$ and with $\mu(\{f^{-1}\{t_j\}\}) = 0$ for each $j = 1, \ldots, N$. Notice that the latter requirement is no problem because $\mu(\{f^{-1}\{t\}\}) = 0$ for all but a countable set of $t > 0$, by virtue of the fact that $\mu\{x \in X : f(x) > 0\} < \infty$.

Now let $U_j = f^{-1}\{(t_{j-1}, t_j)\}$, $j = 1, \ldots, N$. (Notice that then the U_j are pairwise disjoint and each $U_j \subset K$, where K, compact, is the support of f.) Now by the definition (1) we can find $g_j \in \mathcal{K}_+$ such that $g_j \le 1$, support $g_j \subset U_j$, and $\lambda(g_j) \ge \mu(U_j) - \varepsilon/N$. Also for any compact $K_j \subset U_j$ we can construct a function $h_j \in \mathcal{K}_+$ with $h_j \equiv 1$ in a neighborhood of $K_j \cup$ support g_j, support $h_j \subset U_j$, and $h_j \le 1$ everywhere. Then $h_j \ge g_j$, $h_j \le 1$ everywhere and support h_j is a compact subset of U_j and so

$$\mu(U_j) - \varepsilon/N \le \lambda(g_j) \le \lambda(h_j) \le \mu(U_j), \quad j = 1, \ldots, N. \tag{5}$$

Since μ is a Radon measure, we can in fact choose the compact $K_j \subset U_j$ such that $\mu(U_j \setminus K_j) < \varepsilon/N$. Then, because $\{x : (f - f\sum_{j=1}^N h_j)(x) > 0\} \subset \cup(U_j \setminus K_j)$, by (4) we have

$$\lambda(f - f\textstyle\sum_{j=1}^N h_j) \le \varepsilon \sup f. \tag{6}$$

Then by using (5), (6) and the linearity of λ (together with the fact $t_{j-1}h_j \le f h_j \le t_j h_j$) for each $j = 1, \ldots, N$), we see that

$$\begin{aligned}\textstyle\sum_{j=1}^N t_{j-1}\mu(U_j) - \varepsilon \sup f &\le \lambda(f\textstyle\sum_j h_j) \le \lambda(f) \le \lambda(f\sum_j h_j) + \varepsilon \sup f \\ &\le \textstyle\sum_{j=1}^N t_j \mu(U_j) + \varepsilon \sup f.\end{aligned}$$

Since trivially

$$\textstyle\sum_{j=1}^N t_{j-1}\mu(U_j) \le \int_X f\, d\mu \le \sum_{j=1}^N t_j \mu(U_j),$$

we then have

$$\begin{aligned}-\varepsilon(\mu(K) + \sup f) &\le -\textstyle\sum_{j=1}^N (t_j - t_{j-1})\mu(U_j) - \varepsilon \sup f \\ &\le \int_X f\, d\mu - \lambda(f) \\ &\le \textstyle\sum_{j=1}^N (t_j - t_{j-1})\mu(U_j) + \varepsilon \sup f \le \varepsilon(\mu(K) + \sup f),\end{aligned}$$

where $K = \text{support} f$. This completes the proof of 4.12. □

We can now state the Riesz Representation Theorem. In the statement, $C_c(X, H)$ will denote the set of vector functions $f : X \to H$ which are continuous and which have compact support, where H is a given finite dimensional real Hilbert space with inner product $\langle\,,\,\rangle$ and inner product norm $|\,|$.

4.14 Theorem (Riesz Representation Theorem.) *Suppose X is a locally compact Hausdorff space, and $L : C_c(X, H) \to \mathbb{R}$ is linear with*

$$\sup_{f \in C_c(X,H), |f| \le 1, \text{support } f \subset K} L(f) < \infty \text{ whenever } K \subset X \text{ is compact.}$$

Then there is a Radon measure μ on X and μ-measurable $\nu : X \to H$ with $|\nu| = 1$ μ-a.e. on X, and

$$L(f) = \int_X \langle f, \nu \rangle\, d\mu \text{ for any } f \in C_c(X, H).$$

Proof: By using an orthonormal basis for H, it suffices to prove the theorem with $H = \mathbb{R}^n$. We first define

$$\lambda(f) = \sup_{\omega\in C_c(X,\mathbb{R}^n),\,|\omega|\le f} L(\omega)$$

for any $f \in \mathcal{K}_+$. We claim that λ has the linearity properties of the lemma. Indeed it is clear that $\lambda(cf) = c\lambda(f)$ for any constant $c \ge 0$ and any $f \in \mathcal{K}_+$. Now let $f, g \in \mathcal{K}_+$, and notice that if $\omega_1, \omega_2 \in C_c(X,\mathbb{R}^n)$ with $|\omega_1| \le f$ and $|\omega_2| \le g$, then $|\omega_1 + \omega_2| \le f + g$ and hence $\lambda(f+g) \ge L(\omega_1) + L(\omega_2)$. Taking sup over all such ω_1, ω_2 we then have $\lambda(f+g) \ge \lambda(f) + \lambda(g)$. To prove the reverse inequality we let $\omega \in C_c(X,\mathbb{R}^n)$ with $|\omega| \le f + g$, and define

$$\omega_1 = \begin{cases} \frac{f}{f+g}\omega & \text{if } f+g>0 \\ 0 & \text{if } f+g=0, \end{cases} \qquad \omega_2 = \begin{cases} \frac{g}{f+g}\omega & \text{if } f+g>0 \\ 0 & \text{if } f+g=0. \end{cases}$$

Then $\omega_1 + \omega_2 = \omega$, $|\omega_1| \le f$, $|\omega_2| \le g$ and it is readily checked that $\omega_1, \omega_2 \in C_c(X,\mathbb{R}^n)$. Then $L(\omega) = L(\omega_1) + L(\omega_2) \le \lambda(f) + \lambda(g)$, and hence taking sup over all such ω we have $\lambda(f+g) \le \lambda(f) + \lambda(g)$. Therefore we have $\lambda(f+g) = \lambda(f) + \lambda(g)$ as claimed. Thus λ satisfies the conditions of the lemma, hence there is a Radon measure μ on X such that

$$\lambda(f) = \int_X f\,d\mu, \quad f \in \mathcal{K}_+, \quad j = 1,\ldots,n.$$

That is, we have

$$(\ddagger) \qquad \sup_{\omega\in C_c(X,\mathbb{R}^n),\,|\omega|\le f} L(\omega) = \int_X f\,d\mu, \quad f \in \mathcal{K}_+.$$

Thus if $j \in \{1,\ldots,n\}$ we have in particular (since $|fe_j| = |f| \in \mathcal{K}_+$ for any $f \in C_c(X,\mathbb{R})$) that

$$|L(fe_j)| \le \int_X |f|\,d\mu \equiv \|f\|_{L^1(\mu)} \quad \forall f \in C_c(X,\mathbb{R}).$$

Thus $L_j(f) \equiv L(fe_j)$ extends to a bounded linear functional on $L^1(\mu)$, and hence by the Riesz representation theorem for $L^1(\mu)$ we know that there is a bounded μ-measurable function ν_j such that

$$L(fe_j) = \int_X f\nu_j\,d\mu, \quad f \in C_c(X,\mathbb{R}).$$

Since any $f = (f_1,\ldots,f_n)$ can be expressed as $f = \sum_{j=1}^n f_j e_j$, we thus deduce

$$(*) \qquad L(f) = \int_X f\cdot\nu\,d\mu, \quad f \in C_c(X,\mathbb{R}^n),$$

where $\nu = (\nu_1,\ldots,\nu_n)$. Then it only remains to check that $|\nu| = 1$ μ-a.e. To see this, first note that by using the Cauchy-Schwarz inequality in the integral on the right of $(*)$ we have for any $f \in \mathcal{K}_+$ that

$$\text{(i)} \qquad \sup_{|g|\le f,\,g\in C_c(X,\mathbb{R}^n)} |L(g)| \le \int_X f|\nu|\,d\mu.$$

On the other hand, we know (since $C_c(X,\mathbb{R}^n)$ is dense in $L^1(\mu)$), we can find a sequence $g_k \in C_c(X,\mathbb{R}^n)$ such that $\lim \int_X |g_k - \hat\nu| = 0$, where $\hat\nu$ is $|\nu|^{-1}\nu$ at points where $\nu \ne 0$ and $\hat\nu = 0$ at all other points. Then of course $\lim \int_X |\hat g_k - \hat\nu| = 0$ with $|\hat g_k| \le 1$, provided we define $\hat g_k = R(g_k)$, with $R(y) = |y|^{-1}y$ if $|y| > 1$ and $R(y) = y$ if $|y| \le 1$, because $|R(y) - v| \le |y - v|$ for any $y, v \in \mathbb{R}^n$ with $|v| = 1$. Thus we deduce that actually equality holds in (i). On the other hand by $(\ddagger)$ for any $f \in \mathcal{K}_+$ we have that the left side of (i) is $\int_X f\,d\mu$. Thus finally $\int_X f\,d\mu = \int_X f|\nu|\,d\mu$, and this evidently implies $|\nu| = 1$ μ-a.e., again using the density of $C_c(X,\mathbb{R})$ in $L^1(\mu)$. □

4.15 Remark: Notice that L extends as a continuous linear functional on $L^\infty(\mu)$.

Using the Riesz Theorem 4.12 we can deduce the following compactness theorem for Radon measures:

4.16 Theorem (Compactness Theorem for Radon Measures.) *Suppose $\{\mu_k\}$ is a sequence of Radon measures on the locally compact, σ-compact Hausdorff space X with the property $\sup_k \mu_k(K) < \infty$ for each compact $K \subset X$. Then there is a subsequence $\{\mu_{k'}\}$ which converges to a Radon measure μ on X in the sense that*

$$\lim \mu_{k'}(f) = \mu(f) \text{ for each } f \in \mathcal{K}(X),$$

where $\mathcal{K}(X)$ denotes the set of continuous functions $f : X \to \mathbb{R}$ with compact support on X and where we use the notation

$$\mu(f) = \int_X f\,d\mu, \quad f \in \mathcal{K}(X).$$

Proof: Let $K_1, K_2, \ldots$ be an increasing sequence of compact sets with $X = \cup_j K_j$ and let $F_{j,k} : C(K_j) \to \mathbb{R}$ be defined by $F_{j,k}(f) = \int_{K_j} f\,d\mu_k$, $k = 1,2,\ldots$. By the Alaoglu theorem there is a subsequence $F_{j,k'}$ and a non-negative bounded functional $F_j : C(K_j) \to \mathbb{R}$ with $F_{j,k'}(f) \to F_j(f)$ for each $f \in C(K_j)$. By choosing the subsequences successively and taking a diagonal sequence, we then get a subsequence $\mu_{k'}$ and a non-negative linear $F : \mathcal{K}(X) \to \mathbb{R}$ with $\int_X f\,d\mu_{k'} \to F(f)$ for each $f \in \mathcal{K}(X)$, where $F(f) = F_j(f|K_j)$ whenever spt $f \subset K_j$. (Notice that this is unambiguous because if spt $f \subset K_j$ and $\ell > j$ then $F_\ell(f|K_\ell) = F_j(f|K_j)$ by construction.) Then by applying Theorem 4.12 we have a Radon measure μ on X such that $F(f) = \int_X f\, d\mu$ for each $f \in \mathcal{K}(X)$, and so $\int_X f\, d\mu_{k'} \to \int_X f\, d\mu$ for each $f \in \mathcal{K}(X)$. $\square$

Chapter 2

Some Further Preliminaries from Analysis

Here we develop the necessary further analytical background material needed for later developments. In particular we prove some basic results about Lipschitz and BV functions, and we also present the basic facts concerning C^k submanifolds of Euclidean space. There is also a brief treatment of the area and co-area formulae and a discussion of first and second variation formulae for C^2 submanifolds of Euclidean space. These latter topics will be discussed in a much more general context later.

1 Lipschitz Functions

If X is a metric space with metric d, recall that a function $f : X \to \mathbb{R}$ is said to be *Lipschitz* if there is $L < \infty$ such that

1.1 $$|f(x) - f(y)| \le L\, d(x, y) \ \forall x, y \in X.$$

Lip f denotes the least such constant L.

First we have the following basic extension theorem.

1.2 Theorem. *If A is a non-empty subset of X and $f : A \to \mathbb{R}$ is Lipschitz, then $\exists\ \overline{f} : X \to \mathbb{R}$ with $\operatorname{Lip} \overline{f} = \operatorname{Lip} f$, and $f = \overline{f}|A$.*

Proof: With $L = \operatorname{Lip} f$, we claim that

$$\overline{f}(x) = \inf_{z\in A}(f(z) + Ld(x,z)),\ x \in X,$$

has the required properties. Indeed if $x \in A$ then trivially $\overline{f}(x) \le f(x)$ and also $\overline{f}(x)-f(x) = \inf_{z\in A}(f(z)-f(x)+Ld(x,z)) \ge \inf_{z\in A}(-Ld(x,z)+Ld(x,z)) = 0$, so $\overline{f}(x) = f(x)$.

Also for any $x_1, x_2 \in X$

$$\begin{aligned}\overline{f}(x_1) - \overline{f}(x_2) &= \sup_{z_2\in A}\inf_{z_1\in A}\big(f(z_1) + Ld(x_1,z_1) - f(z_2) - Ld(x_2,z_2)\big)\\ &\le \sup_{z_2\in A}\big(Ld(x_1,z_2) - Ld(x_2,z_2)\big) \le Ld(x_1,x_2)\end{aligned}$$

and the reverse inequality holds by interchanging x_1, x_2. □

1.3 Remark: Observe that the above proof has a geometric interpretation: the graph of the extension $\overline{f}$ is obtained by taking the "lower envelope" (inf) of all the half-cones $C_z = \{(x,y) \in X\times\mathbb{R} : y = f(z) + Ld(x,z)\}$; notice that C_z is a half-cone of slope L with vertex on the graph of the original function f.

Next we need the theorem of Rademacher concerning differentiability of Lipschitz functions on $\mathbb{R}^n$. (The proof given here is due to C.B. Morrey.)

1.4 Theorem (Rademacher's theorem.) *If f is Lipschitz on $\mathbb{R}^n$, then f is differentiable $\mathcal{L}^n$-almost everywhere; that is, the gradient $\nabla f(x) = (D_1 f(x),\dots,D_n f(x))$ exists and*

$$(*)\qquad \lim_{y\to x}\frac{f(y)-f(x)-\nabla f(x)\cdot(y-x)}{|y-x|} = 0$$

for $\mathcal{L}^n$-a.e. $x \in \mathbb{R}^n$.

Proof: Let $v \in \mathbb{S}^{n-1}$, and whenever it exists let $D_v f(x)$ denote the directional derivative $\frac{d}{dt}f(x+tv)\big|_{t=0}$. Since $\left|\frac{f(y)-f(x)}{|y-x|}\right| \le \operatorname{Lip} f$ for $y \ne x$ (so $|D_v f| \le \operatorname{Lip} f$ whenever it exists) and we see that $D_v f(x)$ exists precisely when the bounded functions

$$\limsup_{t\to 0}\frac{f(x+tv)-f(x)}{t},\quad \liminf_{t\to 0}\frac{f(x+tv)-f(x)}{t}$$

coincide. Now $\limsup_{t\to 0}\frac{f(x+tv)-f(x)}{t} = \lim_{j\to\infty}\sup_{0<|t|<j^{-1}}\frac{f(x+tv)-f(x)}{t}$ which is Borel measurable because $\sup_{0<|t|<j^{-1}}\frac{f(x+tv)-f(x)}{t}$ is lower semi-continuous, and hence Borel measurable, for each j. Similarly $\liminf_{t\to 0}\frac{f(x+tv)-f(x)}{t}$ is Borel measurable, so the set $A_v = \{x \in \mathbb{R}^n : D_v f \text{ does not exist}\}$ is Borel measurable and hence $\mathcal{L}^n$-measurable. However $\varphi(t) = f(x+tv)$ is an absolutely continuous function of $t \in \mathbb{R}$ for any fixed x and v, and hence is differentiable for almost all t. Thus A_v intersects every line L which is parallel to v in a set of $\mathcal{H}^1$ measure zero and hence by Fubini's theorem the Borel set A_v has $\mathcal{L}^n$-measure zero for each v. That is, for each $v \in \mathbb{S}^{n-1}$,

$$(1)\qquad D_v f(x) \text{ exists } \mathcal{L}^n\text{-a.e. } x \in \mathbb{R}^n.$$

Now take any $C_c^\infty(\mathbb{R}^n)$ function ζ and note that for any $h > 0$

$$(2)\qquad \int_{\mathbb{R}^n}\frac{f(x+hv)-f(x)}{h}\zeta(x)\,d\mathcal{L}^n(x) = -\int_{\mathbb{R}^n}\frac{\zeta(x)-\zeta(x-hv)}{h}f(x)\,d\mathcal{L}^n(x)$$

(by the change of variable $z = x + hv$ in the first part of the integral on the left). Using the dominated convergence theorem and (1) we then have

$$(3)\qquad \begin{aligned}\int D_v f\zeta &= -\int f D_v\zeta = -\int f v\cdot\nabla\zeta\\ &= -\textstyle\sum_{j=1}^n v^j\displaystyle\int f D_j\zeta = +\textstyle\sum_{j=1}^n v^j\displaystyle\int \zeta D_j f = \int \zeta v\cdot\nabla f,\end{aligned}$$

where ∇f is the gradient of f (i.e. $\nabla f = (D_1 f,\dots,D_n f)0$ all integrals are with respect to Lebesgue measure on $\mathbb{R}^n$, and where we have used Fubini's theorem and the absolute continuity of f on lines to justify the integration by parts. Since ζ is arbitrary in (3) we have, for $\mathcal{L}^n$-a.e. $x \in \mathbb{R}^n$,

$$(4)\qquad D_v f(x) = v\cdot\nabla f(x)\ \forall\, v \in S^{n-1}.$$

Of course at such points x we have

$$(5)\qquad |\nabla f(x)| = \sup_{|v|=1} D_v f(x) \le L.$$

Now let $v_1, v_2, \dots$ be a countable dense subset of $\mathbb{S}^{n-1}$, and let

$$A_k = \{x : \nabla f(x), D_{v_k} f(x) \text{ exist and } D_{v_k} f(x) = v_k\cdot\nabla f(x)\}.$$

Then $A = \cap_{k=1}^\infty A_k$ we have by (4) that

$$(6)\qquad \mathcal{L}^n(\mathbb{R}^n\setminus A) = 0,\ D_{v_k} f(x) = v_k\cdot\nabla f(x)\ \forall x \in A,\ k = 1,2,\dots.$$

Using this, we are now going to prove that f is differentiable at each point x of A. To see this, for any $x \in A$, $v \in \mathbb{S}^{n-1}$ and $h > 0$ define

(7) $$Q(x,v,h) = \frac{f(x+hv)-f(x)}{h} - v \cdot \nabla f(x),$$

so by (6)

(8) $$\lim_{h\to 0} Q(x,v_j,h) = 0, \; x \in A, \; j = 1,2,\ldots.$$

Now for any given $\varepsilon > 0$, select P large enough so that

(9) $$S^{n-1} \subset \cup_{i=1}^{P} B_\varepsilon(v_i),$$

and for each $i = 1,\ldots,P$ use (8) to choose $\delta_i > 0$ so that

(10) $$0 < |h| < \delta_j \Rightarrow |Q(x,v_i,h)| < \varepsilon.$$

By (9), for any $v \in S^{n-1}$ we can select $i \in \{1,\ldots,P\}$ with $|v - v_i| < \varepsilon$, and hence by (10)

$$\begin{aligned}|Q(x,h,v)| &\le |Q(x,v,h) - Q(x,v_i,h)| + |Q(x,v_i,h)| \\ &\le |h|^{-1}|f(x+hv) - f(x+hv_i)| + |v - v_i||\nabla f(x)| + |Q(x,v_i,h)| \\ &< (2L+1)\varepsilon \text{ for all } 0 < |h| < \delta = \min\{\delta_1,\ldots,\delta_P\}\end{aligned}$$

by (5). Thus $v \in S^{n-1}$ and $0 < |h| < \delta \Rightarrow |Q(x,h,v)| < (2L+1)\varepsilon$, hence f is differentiable at x. $\square$

We shall need the following C^1 approximation theorem for Lipschitz functions in our discussion of rectifiable sets in the next chapter.

1.5 Theorem. (C^1 Approximation Theorem.) *Suppose $f : \mathbb{R}^n \to \mathbb{R}$ is Lipschitz. Then for each $\varepsilon > 0$ there is a $C^1(\mathbb{R}^n)$ function g with*

$$\mathcal{L}^n\big(\{x : f(x) \neq g(x)\} \cup \{x : \nabla f(x) \neq \nabla g(x)\}\big) < \varepsilon.$$

Before we begin the proof of 1.5 we need to recall Whitney's extension theorem for C^1 functions:

1.6 Theorem (Whitney Extension Theorem.) *If $A \subset \mathbb{R}^n$ is closed and if $h : A \to \mathbb{R}$ and $\nu : A \to \mathbb{R}^n$ are continuous, and if for each compact $K \subset A$*

(‡) $$\lim_{y\to x,\, y\in A} R(x,y) = 0 \;\; \textit{uniformly for } x \in K,$$

where

$$R(x,y) = \frac{h(y) - h(x) - \nu(x)\cdot(y-x)}{|x-y|},$$

then there is a C^1 function $g : \mathbb{R}^n \to \mathbb{R}$ such that $g = h$ and $\nabla g = \nu$ on A.

(For the proof see for example [EG92] or [Fed69]; for the case $n = 1$, see Remark 1.7 (2) below.)

1.7 Remarks: (1) The hypothesis 1.6 (‡) above cannot be weakened to the requirement that

$$\lim_{y\to x,\, y\in A} R(x,y) = 0, \quad x \in A.$$

For instance we have the example (for $n = 1$) when $A = \{0\} \cup (\cup_{k=1}^{\infty}\{1/k\})$ and $h(0) = 0$, $h(1/k) = (-1)^k/k^{3/2}$, $\nu \equiv 0$. Evidently in this case we do have $\lim_{y\to x,\, y\in A} R(x,y) = 0 \;\forall x \in A$, but there is no C^1 extension because

$$\frac{|h(1/k) - h(1/(k+1))|}{(1/k - 1/(k+1))} \to \infty \text{ as } k \to \infty.$$

In fact the condition 1.6 (‡) is *equivalent* to the existence of a C^1 extension g of f with $\nabla g = \nu$ on A. Indeed if g is such an extension and if $K \subset A$ is compact then for $x, y \in K$ we have

$$\begin{aligned}R(x,y) &= f(y) - f(x) - \nu(x)\cdot(y-x) = g(y) - g(x) - \nabla g(x)\cdot(y-x) \\ &= \int_0^1 \tfrac{d}{dt} g(x + t(y-x))\,dt - \nabla g(x)\cdot(y-x) \\ &= \int_0^1 \big(\nabla g(x+t(y-x)) - \nabla g(x)\big)\cdot(y-x)\,dt\end{aligned}$$

and, since ∇g is uniformly continuous on the convex hull of K, we do indeed have 1.6 (‡).

(2) In the case $n = 1$, the Whitney Extension Theorem 1.6 above has a simple direct proof. Namely in this case define

$$R(x,y) = \frac{h(y) - h(x)}{y - x} - \nu(x)$$

and note that the hypothesis 1.6 (‡) guarantees that for each compact subset C of A we have a function ε_C with $\varepsilon_C(t) \downarrow 0$ as $t \downarrow 0$, and

$$|R(x,y)| \le \varepsilon_C\big(|x-y|\big) \;\forall x, y \in C.$$

In particular this implies

(‡) $$|\nu(x) - \nu(y)| \le 2\varepsilon_C\big(|x-y|\big) \;\forall x, y \in C.$$

Also $\mathbb{R} \setminus A$ is a countable disjoint union of open intervals $I_1, I_2, \ldots$. If $I_j = (a,b)$, we then select $g_j \in C^1([a,b])$ as follows

$$g_j(a) = h(a),\; g_j(b) = h(b),\; g_j'(a) = \nu(a),\; g_j'(b) = \nu(b)$$

and

$$\sup_{x \in I_j} |g_j'(x) - v(a)| \leq 3\varepsilon_C(b-a), \; C = [a-1, b+1] \cap A.$$

This is possible by (‡), with $(x,y) = (a,b)$. One now defines $g(x) = g_j(x)$ $\forall x \in I_j$, $j = 1, 2, \ldots$, and $g(x) = h(x)$ $\forall x \in A$. It is then easy to check $g \in C^1(\mathbb{R})$ and $g' = v$ on A.

Proof of Theorem 1.5: By Rademacher's Theorem f is differentiable on a set $B \subset \mathbb{R}^n$ with $\mathcal{L}^n(\mathbb{R}^n \setminus B) = 0$. By Lusin's theorem (which applies to sets of infinite measure for $\mathcal{L}^n$) there is a closed set $C \subset B$ such that $\nabla f | C$ is continuous and $\mathcal{L}^n(\mathbb{R}^n \setminus C) < \varepsilon/2$. On C we define $h(x) = f(x)$, $v(x) = \nabla f(x)$ and $R(x,y)$ for $x, y \in C$ is as defined in 1.6 (‡). Evidently (since $C \subset B$) we have

$$\lim_{y \to x,\, y \in C} R(x,y) = 0 \;\; \forall x \in C,$$

but *not* necessarily uniformly with respect to x on compact subsets of C as in 1.6 (‡). We therefore proceed as follows. For each $k = 1, 2, \ldots$ let

$$\eta_k(x) = \sup\{|R(x,y)| : y \in C \cap (B_{1/k}(x) \setminus \{x\})\}.$$

Then $\eta_k \downarrow 0$ pointwise in C, and hence by *Egoroff's Theorem* (applied to the finite measure sets $C \cap B_j(0)$, $j = 1, 2, \ldots$) there is an $\mathcal{L}^n$-measurable set $A_0 \subset C$ such that $\mathcal{L}^n(C \setminus A_0) < \varepsilon/4$ and η_k converges uniformly to zero on each bounded subset of A_0. Since we can take a closed set $A \subset A_0$ with $\mathcal{L}^n(A_0 \setminus A) < \varepsilon/4$ we thus have $\mathcal{L}^n(\mathbb{R}^n \setminus A) < \varepsilon$ and 1.6 (‡) holds. Hence we can apply the Whitney Theorem 1.6, and the proof is complete. □

Next we establish some basic facts about Hausdorff measure of Lipschitz images:

1.8 Theorem. *Suppose X, Y metric spaces with X σ-compact (i.e. there are compact $K_1, K_2, \ldots$ with $X = \cup_j K_j$), $A \subset X$ with A $\mathcal{H}^m$-measurable and $\mathcal{H}^m(A) < \infty$, $f : A \to Y$ Lipschitz, and let $\mathcal{N}(f,y) = \mathcal{H}^0(f^{-1}y)$ (i.e. $\mathcal{N}(f,y)$ is the multiplicity function, counting the number of points, possibly ∞, in the preimage $f^{-1}y$). Then*

(i) *$f(A)$ is $\mathcal{H}^m$-measurable with $\mathcal{H}^m(f(A)) \leq (\mathrm{Lip}\, f)^m \mathcal{H}^m(A)$,*

(ii) *$\mathcal{N}(f,y)$ is an $\mathcal{H}^m$-measurable function of $y \in Y$ with*

$$\int_Y \mathcal{N}(f,y)\, d\mathcal{H}^m \leq (\mathrm{Lip}\, f)^m \mathcal{H}^m(A).$$

Proof: (i) Observe first that if $\delta > 0$ and if $C_1, C_2, \ldots$ are chosen with $A \subset \cup_j C_j$ and $\operatorname{diam} C_j < \delta$ for each j, then $f(A) \subset \cup_j f(C_j)$ and $\operatorname{diam}(f(C_j)) \leq (\mathrm{Lip\ f})\delta < (1 + \mathrm{Lip}\, f)\delta$. Hence

$$\mathcal{H}^m_{(1+\mathrm{Lip}\, f)\delta}(f(A)) \leq \textstyle\sum_j \omega_m (\operatorname{diam} f(C_j)/2)^m \leq (\mathrm{Lip}\, f)^m \sum_j \omega_m (\operatorname{diam} C_j/2)^m,$$

and taking inf over all such collections $\{C_j\}$ and then letting $\delta \downarrow 0$ we obtain $\mathcal{H}^m(f(A)) \leq (\mathrm{Lip}\, f)^m \mathcal{H}^m(A)$. Notice this part of the argument is valid even if A is not $\mathcal{H}^m$-measurable and did not depend on the σ-compactness of X.

Next suppose that A is $\mathcal{H}^m$-measurable (e.g. A is Borel) and observe that the regularity property 1.15 (2) of Ch. 1 together with the σ-compactness of X implies that we can take a sequence $K_1, K_2, \ldots$ of compact sets in X with $K_j \subset A$ for each j and $\mathcal{H}^m(A \setminus (\cup_j K_j)) = 0$. By the first part above we then have $\mathcal{H}^m(f(A \setminus (\cup_j K_j))) = 0$ and hence $f(A) = f(A \setminus (\cup_j K_j)) \cup (\cup_j f(K_j))$. Now $f(K_j)$ is compact, hence Borel, for each j, and, by the first part of the discussion above, $\mathcal{H}^m(f(A \setminus (\cup_j K_j))) \leq (\mathrm{Lip}\, f)^m \mathcal{H}^m(A \setminus (\cup_j K_j)) = 0$ so $f(A)$ is $\mathcal{H}^m$-measurable as claimed. This completes the proof of (i).

To prove (ii) observe that, by the σ-compactness of X, for each $i = 1, 2, \ldots$ we can partition A into a disjoint union $\cup_{j=1}^\infty A_{ij}$ where each A_{ij} is $\mathcal{H}^m$-measurable and $\operatorname{diam}(A_{ij}) < 1/i$; furthermore we can do this successively, partitioning each A_{ik} to give the new sets A_{i+1j}, so that each of the sets A_{i+1j} is contained in one of the A_{ik}. Observe that then $\sum_j \chi_{f(A_{ij})}$ is a non-negative function which is $\mathcal{H}^m$-measurable by (i) above and which increases pointwise (at *every* point) to $\mathcal{N}(f,y)$, and so $\mathcal{N}(f,y)$ is $\mathcal{H}^m$-measurable and by the monotone convergence theorem

$$\int_Y \mathcal{N}(f,y)\, d\mathcal{H}^m(y) = \lim_{i \to \infty} \int_Y \textstyle\sum_j \chi_{f(A_{ij})}\, d\mathcal{H}^m = \lim_{i \to \infty} \sum_j \mathcal{H}^m(f(A_{ij})),$$

and

$$\textstyle\sum_j \mathcal{H}^m(f(A_{ij})) \leq (\mathrm{Lip}\, f)^m \sum_j \mathcal{H}^m(A_{ij}) = (\mathrm{Lip}\, f)^m \mathcal{H}^m(A)$$

by part (i) above. □

Next we want to extend the inequality of Theorem 1.8 (ii) to the case when the k-dimensional Hausdorff measure of $f^{-1}y$ (instead of $\mathcal{H}^0(f^{-1}y)$) appears on the left. For this we assume for convenience that $Y = \mathbb{R}^m$ (more general cases, e.g. when Y is a metric space such that each closed ball is compact, are discussed in [Fed69, 10.2.25], but the case $Y = \mathbb{R}^m$ is adequate for the subsequent development here, and furthermore the proof is relatively elementary in this case).

1.9 Theorem. *Suppose X is a σ-compact metric space, $m \in \{1, 2, \ldots\}$, $k > 0$ (k need not be an integer), $A \subset X$ is $\mathcal{H}^{m+k}$-measurable and $\mathcal{H}^{m+k}(A) < \infty$, and $f : A \to \mathbb{R}^m$ is Lipschitz. Then $\mathcal{H}^k(f^{-1}y)$ is an $\mathcal{L}^m$-measurable function of $y \in \mathbb{R}^m$ and*

$$\int_{\mathbb{R}^m} \mathcal{H}^k(f^{-1}y)\, d\mathcal{L}^m(y) \leq \frac{\omega_m \omega_k}{\omega_{m+k}} (\mathrm{Lip}\, f)^m \mathcal{H}^{m+k}(A).$$

1.10 Remark: At one step in the proof below we are going to use the *upper Lebesgue integral* $\int^*_{\mathbb{R}^m} f\, d\mathcal{L}^m$ of a not necessarily measurable function $f : \mathbb{R}^m \to [0, \infty]$. This is defined by

$$\int_{\mathbb{R}^m}^* f\, d\mathcal{L}^m = \inf_{\psi \geq f, \psi \text{ measurable}} \int_{\mathbb{R}^m} \psi\, d\mathcal{L}^m.$$

Observe that then there is always a measurable function ψ_f which attains the inf; that is, $\psi_f \geq f$ and

$$\int_{\mathbb{R}^m}^* f\, d\mathcal{L}^m = \int_{\mathbb{R}^m} \psi_f\, d\mathcal{L}^m,$$

and if $\int_{\mathbb{R}^m}^* f\, d\mathcal{L}^m < \infty$ the function ψ_f is unique up to change on a set of measure zero. Notice also that if $\{f_i\}$ is an increasing sequence of maps $\mathbb{R}^m \to [0,\infty]$ and if $f = \lim_{i\to\infty} f_i$, then $\lim_{i\to\infty} \int_{\mathbb{R}^m}^* f_i\, d\mathcal{L}^m = \int_{\mathbb{R}^m}^* f\, d\mathcal{L}^m$.

Proof of 1.9: For each $i = 1, 2, \ldots$ pick closed subsets $C_{i1}, C_{i2}, \ldots$ with $\operatorname{diam} C_{ij} < 1/i$ $A \subset \cup_j C_{ij}$ and

(1) $$\textstyle\sum_j \omega_{m+k} (\operatorname{diam} C_{ij}/2)^{m+k} \leq \mathcal{H}^{m+k}_{1/i}(A) + 1/i,$$

and let g_i be the Borel measurable simple functions defined by

$$g_i = \textstyle\sum_j \omega_k (\operatorname{diam} C_{ij}/2)^k \chi_{f(C_{ij})}.$$

Notice that, by σ-compactness of X, each C_{ij} is a countable union of compact sets and hence indeed $f(C_{ij})$ is a Borel set and g_i is Borel measurable as claimed. Also notice that

(2) $$\begin{aligned} g_i(y) = \textstyle\sum_j \omega_k (\operatorname{diam} C_{ij}/2)^k \chi_{f(C_{ij})}(y) &= \textstyle\sum_{j: f^{-1}y \cap C_{ij} \neq \emptyset} \omega_k (\operatorname{diam} C_{ij}/2)^k \\ &\geq \mathcal{H}^k_{1/i}(A \cap f^{-1}y), \end{aligned}$$

and hence by integrating each side with respect to $\mathcal{H}^m$-measure on $\mathbb{R}^m$ we then have, by (1) and (2),

$$\begin{aligned} \int_{\mathbb{R}^m}^* \mathcal{H}^k_{1/i}(A \cap f^{-1}y)\, d\mathcal{L}^m(y) &\leq \int_{\mathbb{R}^m} g_i(y)\, d\mathcal{L}^m(y) \\ &= \textstyle\sum_j \omega_k (\operatorname{diam} C_{ij}/2)^k \mathcal{L}^m(f(C_{ij})) \leq \sum_j \omega_k \omega_m (\operatorname{Lip} f)^m (\operatorname{diam} C_{ij}/2)^{m+k} \\ &\leq \Big(\frac{\omega_m \omega_k}{\omega_{m+k}}\Big) (\operatorname{Lip} f)^m \big(\mathcal{H}^{m+k}_{1/i}(A) + 1/i\big), \end{aligned}$$

where the notation $\int^*$ is as in Remark 1.10 above and where we used $\mathcal{L}^m(f(C_{ij})) \leq \omega_m \big(\frac{\operatorname{diam} f(C_{ij})}{2}\big)^m$ (by the isodiametric inequality 2.10) $\leq \omega_m (\operatorname{Lip} f)^m \big(\frac{\operatorname{diam} C_{ij}}{2}\big)^m$. Letting $i \to \infty$, we conclude

$$\int_{\mathbb{R}^m}^* \mathcal{H}^k(A \cap f^{-1}y)\, d\mathcal{L}^m(y) \leq \Big(\frac{\omega_m \omega_k}{\omega_{m+k}}\Big) (\operatorname{Lip} f)^m \mathcal{H}^{m+k}(A).$$

It remains to check that $\mathcal{H}^k(A \cap f^{-1}y)$ is an $\mathcal{H}^m$-measurable function of $y \in \mathbb{R}^m$ (which will enable us to replace the upper integral on the left of the above inequality with the standard integral). This is left as an exercise (Q.5 of hw2). □

We conclude this section with a discussion of Lipschitz domains in $\mathbb{R}^n$.

1.11 Definition: A bounded open set $\Omega \subset \mathbb{R}^n$ is said to be a Lipschitz domain if there are constants $0 < \sigma \leq \tau$ such that $\forall\, y \in \partial\Omega$ there is a $v \in \mathbb{S}^{n-1}$ and a Lipschitz function $u : B_\sigma(0) \cap v^\perp \to (-\tau, \tau)$ such that

$$\begin{aligned} U_y \cap \Omega &= \{y + x + tv : x \in \breve{B}_\sigma(0) \cap v^\perp,\ t < u(x)\} \\ U_y \cap \partial\Omega &= \{y + x + tv : x \in \breve{B}_\sigma(0) \cap v^\perp,\ t = u(x)\}, \end{aligned}$$

where U_y is the open neighborhood of y given by

$$U_y = \{y + x + tv : x \in \breve{B}_\sigma(0) \cap v^\perp,\ -\tau < t < \tau\}.$$

Thus, roughly speaking, Ω is Lipschitz means that locally, near each of its points, $\partial\Omega$ can be expressed as the graph of a Lipschitz function.

Of course the bounded open convex subsets of $\mathbb{R}^n$ are automatically Lipschitz domains; more precisely, we have the following lemma:

1.12 Lemma. *Suppose that $\Omega \subset \mathbb{R}^n$ is an open and convex with $0 \in \Omega$, and let $R > 0, \delta \in (0,1)$ be such that $B_{\delta R}(0) \subset \Omega \subset B_R(0)$. Then Ω is Lipschitz; in fact for each $y \in \partial\Omega$ there is a Lipschitz function*

$$u : \breve{B}_{\delta/2}(0) \cap y^\perp \to (0,\infty) \text{ with } u(0) \in (\delta, 1],\ \operatorname{Lip} u \leq 2/\delta,$$

and

$$\begin{aligned} U_y^+ \cap \Omega &= \{x + ty : x \in \breve{B}_{\delta R/2}(0) \cap y^\perp,\ 0 < t < u(x)\} \\ U_y^+ \cap \partial\Omega &= \{x + ty : x \in \breve{B}_{\delta R/2}(0) \cap y^\perp,\ t = u(x)\}, \end{aligned}$$

where U_y^+ is the open neighborhood of y defined by

$$U_y^+ = \{x + ty : x \in \breve{B}_{\delta R/2}(0) \cap y^\perp,\ t > 0\}.$$

Proof: By scaling we can assume without loss of generality that $R = 1$, so $B_\delta(0) \subset \Omega \subset B_1(0)$. Let $y \in \partial\Omega$. By applying a suitable rotation we can also assume that $y = \rho e_n$ with $\rho \in (\delta, 1]$. If $p : \mathbb{R}^n = \mathbb{R}^{n-1} \times \mathbb{R} \to \mathbb{R}^{n-1}$ is the projection $(x,t) \mapsto x$ and if $U = \breve{B}^{n-1}_{\delta/2}(0) \times (0,\infty)$ then evidently

(1) $$p(U \cap \partial\Omega) = \breve{B}^{n-1}_{\delta/2}(0).$$

Let $(x_1, t_1), (x_2, t_2) \in U \cap \partial\Omega$ be arbitrary with $t_2 \geq t_1$, and let π be a supporting hyperplane for $\overline{\Omega}$ at (x_1, t_1), so that there is an open half space H with

$$\pi = \partial H,\ B_\delta(0) \subset \Omega \subset H,\ (x_1, t_1) \in \pi.$$

Then $\pi \cap B_\delta(0) = \emptyset$, so π is not a vertical hyperplane and we can write

$$\pi = \{(x,t) : t = t_1 + a \cdot (x - x_1)\} \text{ and } H = \{(x,t) : t < t_1 + a \cdot (x - x_1)\},$$

where $a \in \mathbb{R}^{n-1}$. We must also then have $|a| \le 2/\delta$, since otherwise there is a point $x \in \breve{B}_\delta^{n-1}(0)$ with $a \cdot (x - x_1) = -t_1$ which would imply $(x, 0) \in \pi \cap \breve{B}_\delta(0)$, contradicting $\pi \cap B_\delta(0) = \emptyset$.

Finally $(x_2, t_2) \in \overline{H}$, so $0 \le t_2 - t_1 \le a \cdot (x_2 - x_1)$ and hence

(2) $$0 \le t_2 - t_1 \le 2\delta^{-1}|x_2 - x_1|.$$

The existence of $u : \breve{B}_{\delta/2}^{n-1}(0) \to (0, \infty)$ with $\operatorname{Lip} u \le 2/\delta$ and $\breve{B}_{\delta/2}^{n-1}(0) \times (0, \infty) \cap \partial\Omega = \operatorname{graph} u$ is now a direct consequence of (1),(2). □

2 BV Functions

In this section we gather together the basic facts about locally BV (i.e. bounded variation) functions which will be needed later.

First recall that if U is open in $\mathbb{R}^n$ and if $u \in L^1_{\text{loc}}(U)$, then u is said to be in $BV_{\text{loc}}(U)$ if for each $W \subset\subset U$ there is a constant $c(W) < \infty$ such that

2.1 $$\int_W u \operatorname{div} g \, d\mathcal{L}^n \le c(W) \sup |g|$$

for all vector functions $g = (g^1, \dots, g^n)$, $g^j \in C_c^\infty(W)$. Notice that this means that the functional $\int_U u \operatorname{div} g$ extends uniquely to give a (real-valued) linear functional on $\mathcal{K}(U, \mathbb{R}^n) \equiv \{\text{continuous } g = (g^1, \dots, g^n) : U \to \mathbb{R}^n \text{ with } \operatorname{spt}|g| \text{ compact}\}$ which is bounded on

$$\mathcal{K}_W(U, \mathbb{R}^n) \equiv \{g \in \mathcal{K}(U, \mathbb{R}^n) : \operatorname{spt}|g| \subset W\}$$

for every $W \subset\subset U$. Then, by the Riesz Representation Theorem 4.14 of Ch. 1, there is a Radon measure μ on U and a μ-measurable function $\nu = (\nu^1, \dots, \nu^n)$, $|\nu| = 1$ a.e., such that

2.2 $$\int_U u \operatorname{div} g \, d\mathcal{L}^n = \int_U g \cdot \nu \, d\mu.$$

Thus, in the language of distribution theory, the generalized derivatives $D_j u$ of u are represented by the *signed measures* $\nu_j \, d\mu$, $j = 1, \dots, n$. For this reason we often denote the total variation measure μ of Ch. 1) by $|Du|$. In fact if $u \in W^{1,1}_{\text{loc}}(U)$ we evidently do have $d\mu = |Du| \, d\mathcal{L}^n$ and

2.3 $$\nu_j = \begin{cases} \frac{D_j u}{|Du|} & \text{if } |Du| \ne 0 \\ 0 & \text{if } |Du| = 0. \end{cases}$$

Thus for $u \in BV_{\text{loc}}(U)$, $|Du|$ will henceforth denote the Radon measure on U which is uniquely characterized by

2.4 $$|Du|(W) = \sup_{|g| \le 1,\, \operatorname{spt}|g| \subset\subset W,\, g \text{ Lipschitz}} \int u \operatorname{div} g \, d\mathcal{L}^n, \quad W \text{ open} \subset U.$$

The left side here is more usually denoted $\int_W |Du|$. Indeed if f is any non-negative Borel measurable function function on U, then $\int f \, d|Du|$ is more usually denoted simply by $\int f|Du|$ ($\equiv \int f|Du| \, d\mathcal{L}^n$ in case $u \in W^{1,1}_{\text{loc}}(U)$). We shall henceforth adopt this notation.

There are a number of important results about BV functions which can be obtained by mollification. We let $\varphi_\sigma(x) = \sigma^{-n}\varphi(x/\sigma)$, where φ is a symmetric mollifier (so that $\varphi \in C_c^\infty(\mathbb{R}^n)$, $\varphi \ge 0$, $\operatorname{spt}\varphi \subset B_1(0)$, $\int_{\mathbb{R}^n} \varphi = 1$, and $\varphi(x) = \varphi(-x)$), and for $u \in L^1_{\text{loc}}(U)$ let $u^{(\sigma)} = \varphi_\sigma * \tilde{u}$ be the mollified functions, where we set $\tilde{u} = u$ on U_σ, $\tilde{u} = 0$ outside U_σ, $U_\sigma = \{x \in U : \operatorname{dist}(x, \partial U) > \sigma\}$. A key result concerning mollification is then as follows:

2.5 Lemma. *If $u \in BV_{\text{loc}}(U)$, then $u^{(\sigma)} \to u$ in $L^1_{\text{loc}}(U)$ and $|Du^{(\sigma)}| \to |Du|$ in the sense of Radon measures in U (see* 4.16 *of* Ch. 1*) as $\sigma \downarrow 0$.*

The convergence of $u^{(\sigma)}$ to u in $L^1_{\text{loc}}(U)$ is standard. Thus it remains to prove

(1) $$\lim_{\sigma \downarrow 0} \int f|Du^{(\sigma)}| = \int f|Du|$$

for each $f \in C_c^0(U)$, $f \ge 0$. In fact by definition of $|Du|$ it is rather easy to prove that

(2) $$\int f|Du| \le \liminf_{\sigma \downarrow 0} \int f|Du^{(\sigma)}|,$$

so we only have to check

(3) $$\liminf_{\sigma \downarrow 0} \int f|Du^{(\sigma)}| \le \int f|Du|$$

for each $f \in C_c^0(U)$, $f \ge 0$.

This is achieved as follows: First note that

(4) $$\int f|Du^{(\sigma)}| = \sup_{|g| \le f,\, g \text{ smooth}} \int g \cdot \nabla u^{(\sigma)} \, d\mathcal{L}^n.$$

On the other hand for fixed g with g smooth and $|g| \le f$, and for $\sigma < \operatorname{dist}\{\operatorname{spt} f, \partial U\}$, we have

$$\begin{aligned}\int g \cdot \nabla u^{(\sigma)}\, d\mathcal{L}^n &= -\int u^{(\sigma)} \operatorname{div} g\, d\mathcal{L}^n \\ &= -\int \varphi_\sigma * u \operatorname{div} g\, d\mathcal{L}^n \\ &= -\int u(\varphi_\sigma * \operatorname{div} g)\, d\mathcal{L}^n \\ &= -\int u \operatorname{div}(\varphi_\sigma * g)\, d\mathcal{L}^n.\end{aligned}$$

On the other hand by definition of $|Du|$, the right side here is

$$\leq \int_{W_\sigma} (f + \varepsilon(\sigma))|Du|$$

where $\varepsilon(\sigma) \downarrow 0$, where $W = \operatorname{spt} f$, $W_\sigma = \{x \in U : \operatorname{dist} x, W < \sigma\}$, because

$$\begin{aligned}|\varphi_\sigma * g| &\equiv |(\varphi_\sigma * g^1, \ldots, \varphi_\sigma * g^n)| \\ &\leq \varphi_\sigma * |g| \leq \varphi_\sigma * f\end{aligned}$$

and because $\varphi_\sigma * f \to f$ uniformly in W_{σ_0} as $\sigma \downarrow 0$, where $\sigma_0 < \operatorname{dist}(W, \partial U)$. Thus (3) follows from (4). □

2.6 Theorem (Compactness Theorem for BV Functions.) *If $\{u_k\}$ is a sequence of $BV_{\text{loc}}(U)$ functions satisfying*

$$\sup_{k\geq 1}(\|u_k\|_{L^1(W)} + \int_W |Du_k|) < \infty$$

for each $W \subset\subset U$, then there is a subsequence $\{u_{k'}\} \subset \{u_k\}$ and a $BV_{\text{loc}}(U)$ function u such that $u_{k'} \to u$ in $L^1_{\text{loc}}(U)$ and

$$\int_W |Du| \leq \liminf \int_W |Du_{k'}| \ \forall W \subset\subset U.$$

Proof: By virtue of the previous lemma, in order to prove $u_{k'} \to u$ in $L^1_{\text{loc}}(U)$ for some subsequence $\{u_{k'}\}$, it is enough to prove that the sets

$$\{u \in C^\infty(U) : \int_W (|u| + |Du|)\, d\mathcal{L}^n \leq c(W)\}, \quad W \subset\subset U,$$

(for given constants $c(W) < \infty$) are precompact in $L^1_{\text{loc}}(U)$. For the simple proof of this (involving mollification and Arzela's theorem) see for example [GT01, Theorem 7.22].

Finally the fact that $\int_W |Du| \leq \liminf \int_W |Du_{k'}|$ is a direct consequence of the definition of $|Du|$, $|Du_{k'}|$. □

Next we have the Poincaré inequality for BV functions.

2.7 Lemma. *Suppose U is bounded, open and convex, let $\delta \in (0,1)$ be such that there is $R > 0$ and $\xi \in U$ with $B_{\delta R}(\xi) \subset U \subset B_R(\xi)$, and let $u \in BV(U)$. Then for any $\theta \in (0,1)$ and any $\beta \in \mathbb{R}$ with*

$$(*) \qquad \min\{\mathcal{L}^n\{x \in U : u(x) \geq \beta\}, \mathcal{L}^n\{x \in U : u(x) \leq \beta\}\} \geq \theta\mathcal{L}^n(U).$$

we have

$$\int_U |u - \beta|\, d\mathcal{L}^n \leq CR \int_U |Du|,$$

where $C = C(\theta, \delta, n)$.

Proof: By rescaling $x \mapsto R^{-1}(x - \xi)$ we can without loss of generality assume $R = 1$ and $\xi = 0$.

Let β, θ be as in 2.7 $(*)$ and choose convex $W \subset U$ such that

$$(\ddagger) \qquad \int_W |u - \beta|\, d\mathcal{L}^n \geq \tfrac{1}{2}\int_U |u - \beta|\, d\mathcal{L}^n$$

and such that 2.7 $(*)$ holds with W in place of U and $\theta/2$ in place θ. (For example we may take $W = \{x \in U : \operatorname{dist}(x, \partial U) > \eta\}$ with η small.)

Letting u_σ denote the mollified functions corresponding to u, note that for sufficiently small σ we must have 2.7 $(*)$ with u_σ in place of u, $\theta/4$ in place of θ, and W in place of U. Hence by the usual Poincaré inequality for smooth functions (see e.g. [GT01]) we have, with suitable $\beta^{(\sigma)} \to \beta$ in place of β,

$$\int_W |u_\sigma - \beta^{(\sigma)}|\, d\mathcal{L}^n \leq c \int_W |Du_\sigma|\, d\mathcal{L}^n,$$

$c = c(n, \theta, \delta)$, for all sufficiently small σ. The required inequality now follows by letting $\sigma \downarrow 0$ and using $(\ddagger)$ above together with 2.5. □

2.8 Lemma. *Suppose U, δ, ξ, R are as in 2.7, $u \in BV(\mathbb{R}^n)$ with $\operatorname{spt} u \subset \overline{U}$. Then*

$$\int_{\mathbb{R}^n} |Du| \left(= \int_{\overline{U}} |Du|\right) \leq C\left(\int_U |Du| + R^{-1}\int_U |u|\, d\mathcal{L}^n\right),$$

where $C = C(\delta, n)$.

2.9 Remark: Note that by combining this with the Poincaré inequality in 2.7, we conclude

$$R^{-1}\int_{\mathbb{R}^n} |u - \beta\chi_U| + \int_{\mathbb{R}^n} |D(u - \beta\chi_U)| \leq C\int_U |Du|,$$

$C = C(\theta, \delta)$, whenever β is as in 2.7 $(*)$.

Proof of 2.8: As in the proof of 2.7, we can assume without loss of generality that $R = 1$ and $\xi = 0$.

Let d be the distance function $d(x) = \operatorname{dist}(x, \partial U)$. First note that since U is convex, we must have $U_\sigma = \{x : \operatorname{dist}(x, \partial U) > \sigma\}$ is convex for each $\sigma \geq 0$. Indeed if $\xi \in \partial U_\sigma$, if y is any point of ∂U with $|\xi - y| = \sigma$, and if π is a supporting hyperplane for $\overline{U}$ at y, so that $y \in \pi$ and there is a half-space H with $\partial H = \pi$ and $\overline{U} \subset \overline{H}$, then $y - \xi$ is normal to ∂H and, with H_σ defined to be the half space $\{x - (y - \xi) : x \in H\}$, we have $\overline{U}_\sigma \subset \overline{H}_\sigma$ and $\xi \in \partial H_\sigma$, so U_σ is convex as claimed. Then $0 < d(x) < d(y) \Rightarrow d(x + t(y - x)) \geq d(x)\ \forall t \in [0,1]$ (otherwise $\min_{t\in[0,1]} d(x+t(y-x)) < \min\{d(x), d(y)\}$ which contracts the convexity of U_α, where $\alpha = \min_{t\in[0,1]} d(x + t(y - x))$). Thus $(y - x) \cdot Dd(x) = \frac{d}{dt} d(x + t(y - x))|_{t=0} \geq 0$ for all $x \in U$ such that d is differentiable at x and $d(x) < d(y)$. In particular since $B_\delta(0) \subset U$ (recall we assume $B_{\delta R}(\xi) \subset U \subset B_R(\xi)$ with $R = 1$ and $\xi = 0$) and $|Dd(x)| = 1$ at all points $x \in U$ where d is differentiable, we can take $y = -\delta Dd(x)$, hence $(-x - \delta Dd(x)) \cdot Dd(x) \geq 0$, and so

$$(1) \qquad -x \cdot Dd(x) \geq \delta, \text{ a.e. } x \in \overline{U} \text{ with } d(x) < \operatorname{dist}(B_\delta(0), \partial U).$$

Then we let $\gamma_\sigma : \mathbb{R} \to [0,1]$ be an increasing C^1 function with $\gamma_\sigma(t) \equiv 0$ for $t \leq \sigma/2$ and $\gamma(t) \equiv 1$ for $t \geq \sigma$, and set

$$(2) \qquad \varphi_\sigma = \gamma_\sigma \circ d$$

Then by (1) and (2) we have, for $\sigma < \operatorname{dist}(B_\delta(0), \partial U)$,

$$(3) \qquad \delta\big|D\varphi_\sigma(x)\big| \leq -x \cdot D\varphi_\sigma(x), \quad x \in \overline{U}.$$

Now by definition of $|Dw|$ for $BV_{\text{loc}}(\mathbb{R}^n)$ functions w, we have

$$(4) \qquad \int_{\mathbb{R}^n} |D(\varphi_\sigma u)| \leq \int_{\mathbb{R}^n} |D\varphi_\sigma|\,|u|\,d\mathcal{L}^n + \int_{\mathbb{R}^n} \varphi_\sigma |Du|$$

and by (3)

$$\begin{aligned}(5) \quad \delta\int_{\mathbb{R}^n} |D\varphi_\sigma|\,|u|\,d\mathcal{L}^n &\leq -\int x \cdot D\varphi_\sigma |u|\,d\mathcal{L}^n \\ &= -\int |u| \operatorname{div}(x\varphi_\sigma)\,d\mathcal{L}^n + n\int |u|\varphi_\sigma\,d\mathcal{L}^n \\ &\leq \int_U |D|u|| + n\int_{\mathbb{R}^n} |u|\,d\mathcal{L}^n \text{ (by definition of } |D|u||) \\ &\leq \int_U |Du| + n\int_{\mathbb{R}^n} |u|\,d\mathcal{L}^n\end{aligned}$$

(because $|D|u|| \leq |Du|$ by virtue of 2.5 and the fact that $|D|u|| \leq \liminf_{\sigma\downarrow 0} |D|u_\sigma||$).

Finally, to complete the proof of 2.8, we note that (using the definition of $|Dw|$ for the $BV_{\text{loc}}(\mathbb{R}^n)$ functions $w = u$, $\varphi_\sigma u$, together with the fact that $\varphi_\sigma u \to u$ in $L^1(\mathbb{R}^n)$)

$$\int_{\mathbb{R}^n} |Du| \leq \liminf_{\sigma\downarrow 0} \int_{\mathbb{R}^n} |D(\varphi_\sigma u)|.$$

Then 2.8 follows from (4), (5). □

3 The Area Formula

Recall that if λ is a linear map $\mathbb{R}^n \to \mathbb{R}^n$ and $A \subset \mathbb{R}^n$, then $\mathcal{L}^n(\lambda(A)) = |\det\lambda|\mathcal{L}^n(A)$. More generally if $\lambda : \mathbb{R}^n \to \mathbb{R}^m$, $m = n + k$ with $k \geq 0$, then $\lambda(\mathbb{R}^n) \subset F$ where F is a n-dimensional subspace of $\mathbb{R}^{n+k}$, and hence choosing an orthogonal transformation q of $\mathbb{R}^{n+k}$ such that $q(F) = \mathbb{R}^n \times \{0\}$ and letting $p(x, y) = x$ for $(x, y) \in \mathbb{R}^n \times \mathbb{R}^k = \mathbb{R}^m$, we see that $p\circ q\circ\lambda : \mathbb{R}^n \to \mathbb{R}^n$ and hence $\mathcal{L}^n((p\circ q\circ\lambda)(A)) = |\det(p \circ q \circ \lambda)|\mathcal{L}^n(A)$ for $A \subset \mathbb{R}^n$. Letting $\ell^* : F_2 \to F_1$ denote the adjoint of a linear map ℓ between subspaces F_1, F_2, we then have $(p \circ q \circ \lambda)^*(p \circ q \circ \lambda) = \lambda^* \circ q^* \circ (p^* \circ p) \circ q \circ \lambda$, and since $p^* \circ p$ is the identity on $\mathbb{R}^n \times \{0\}$ this is $\lambda^* \circ \lambda$. Thus $|\det(p \circ q \circ \lambda)| = \sqrt{\det \lambda^* \circ \lambda}$, and since q is an isometry of $\mathbb{R}^m$ we also have $\mathcal{H}^n(q(B)) = \mathcal{H}^n(B)$ for any $B \subset \mathbb{R}^m$, and so finally we obtain the area formula

$$\textbf{3.1} \qquad \mathcal{H}^n(\lambda(A)) = \sqrt{\det(\lambda^* \circ \lambda)}\,\mathcal{H}^n(A), \quad A \subset \mathbb{R}^n,$$

whenever λ is a linear map $\mathbb{R}^n \to \mathbb{R}^{n+k}$, $k \geq 0$.

More generally given a 1:1, Lipschitz map $f : A \to \mathbb{R}^m$ ($A \subset \mathbb{R}^n$ Lebesgue measurable and $m \geq n$) we have, by an approximation argument based on the linear case 3.1 (see [Har79] or [Fed69] for details) that

$$\textbf{3.2} \qquad \mathcal{H}^n(f(A)) = \int_A J_f\,d\mathcal{H}^n.$$

where J_f is the Jacobian of f (or area magnification factor of f) defined by

$$\textbf{3.3} \qquad J_f(y) = \sqrt{\det(df_y)^* \circ (df_y)}$$

Here $df_y : \mathbb{R}^n \to \mathbb{R}^m$ is the induced linear map described in §4, so df_y is represented by the $m\times n$ matrix $\mathcal{J}_f = (D_j f_i)_{i=1,\ldots,m,\ j=1,\ldots,n}$ and $(df_y)^* : \mathbb{R}^{n+\ell} \to T_y M$ denotes the adjoint transformation, represented by the $n\times m$ matrix $(D_j f_i)_{i=1,\ldots,n+\ell, j=1,\ldots,n}$. Observe that in fact then

$$J_f(y) = \sqrt{\det\big(D_i f(y) \cdot D_j f(y)\big)}.$$

If f is not 1:1 we have the general area formula (which actually follows quite easily from the formula 3.2)

3.4
$$\int_{\mathbb{R}^m} \mathcal{H}^0(f^{-1}(y)\cap A)\,d\mathcal{H}^m(y) = \int_A J_f\,d\mathcal{L}^n,$$

where $\mathcal{H}^0$ is 0-dimensional Hausdorff measure i.e. "counting measure," so the term $\mathcal{H}^0(f^{-1}(y)\cap A)$ is the multiplicity function $\mathcal{N}(f,y)$ as in Theorem 1.8 of the present chapter.

More generally still, if h is a non-negative $\mathcal{H}^n$-measurable function on A, then

3.5
$$\int_{\mathbb{R}^{n+\ell}} \textstyle\sum_{x\in f^{-1}(y)} h(x)\,d\mathcal{H}^n(y) = \int_A h\,J_f\,d\mathcal{H}^n.$$

This follows directly from 3.4 if we approximate h by simple functions.

3.6 Examples: (1) *Space curves*: Using the above area formula we first check that $\mathcal{H}^1$-measure agrees with the usual arc-length measure for C^1 curves in $\mathbb{R}^n$. In fact if $\gamma : [a,b] \to \mathbb{R}^n$ is a 1:1 C^1 map then the Jacobian is just $\sqrt{|\dot\gamma|^2} = |\dot\gamma|$, so that 3.2 gives

$$\mathcal{H}^1(\gamma(A)) = \int_A |\dot\gamma|\,d\mathcal{L}^1$$

as required.

(2) *Submanifolds of* $\mathbb{R}^{n+\ell}$: If M is any n-dimensional C^1 manifold of $\mathbb{R}^{n+\ell}$ (see next section for a systematic discussion of submanifolds of $\mathbb{R}^{n+\ell}$), we want to check that $\mathcal{H}^n \llcorner M$ (where $\mathcal{H}^n$ is n-dimensional Hausdorff measure in $\mathbb{R}^{n+\ell}$) agrees with the usual n-dimensional volume measure on M, i.e. that if vol denotes the volume measure (in the usual sense of Riemannian geometry) on the submanifold M, and if $\mathcal{H}^n$ is Hausdorff measure on the ambient space $\mathbb{R}^{n+\ell}$, then for Borel sets $A \subset M$ (or more generally for $\mathcal{H}^n$-measurable sets $A \subset M$) we have

(‡)
$$\mathrm{vol}(A) = \mathcal{H}^n(A).$$

It is enough to check this in a region where a local coordinate representation (see the discussion in §4 below) applies, because we can decompose the Borel set A into a countable pairwise disjoint union of Borel sets A_j, each of which is contained in the image of a local coordinate chart. Thus we suppose U is open in $\mathbb{R}^{n+\ell}$ and that there is a local representation ψ for M such that

$$\psi : W \to \mathbb{R}^{n+\ell} \text{ is } C^1,\ \psi(W) = M\cap U \text{ and } A \subset M \cap U,$$

and let $A_0 = \psi^{-1}(A) \subset W$ be the preimage (of course A_0 is then also Borel). By the area formula

$$\mathcal{H}^n(A) = \int_{A_0} J_\psi\,d\mathcal{L}^n,$$

where $J_\psi = \sqrt{\det(D_i\psi\cdot D_j\psi)}$. Now notice on the other hand that $g_{ij} = D_i\psi\cdot D_j\psi$ is the metric for M (relative to the local coordinates in W) in the usual sense of Riemannian geometry, so this says $\mathcal{H}^n(A) = \int_{A_0}\sqrt{g}\,d\mathcal{L}^n$, where $g = \det(g_{ij})$, and the right side here is indeed the usual definition of vol(A) in the sense of Riemannian geometry, so (‡) is established.

(3) *n-dimensional graphs in* $\mathbb{R}^{n+1}$: If Ω is a domain in $\mathbb{R}^n$ and if $M = \operatorname{graph} u$, where $u \in C^1(\Omega)$, then M is globally represented by the map $\psi : x \mapsto (x, u(x))$; in this case

$$J_\psi(x) \equiv \sqrt{\det\Big(\frac{\partial\psi}{\partial x^i}\cdot\frac{\partial\psi}{\partial x^j}\Big)} \equiv \sqrt{\det(\delta_{ij} + D_iu\,D_ju)} = \sqrt{1+|Du|^2},$$

so $\mathcal{H}^n(M) = \int_\Omega \sqrt{1+|Du|^2}\,dx$ (by (2) above).

4 Submanifolds of $\mathbf{R}^{n+\ell}$

Let M denote an n-dimensional C^r submanifold of $\mathbb{R}^{n+\ell}$, $0 \le \ell$, $r \ge 1$. By this we mean M is a subset of $\mathbb{R}^{n+\ell}$ such that for each $y \in M$ there are open sets $V \subset \mathbb{R}^n$, $W \subset \mathbb{R}^{n+\ell}$ with $y \in W$, and a 1:1 C^r map $\psi : V \to W$ with

4.1
$$\psi(V) = W \cap M$$

and such that ψ is proper (i.e. $K \subset W$ compact $\Rightarrow \psi^{-1}K$ is compact in V) and $D\psi(x)$ has rank n (i.e. maximal rank) at each point $x \in V$. The condition that ψ is proper includes examples like $M = \{(x, \sin 1/x : x > 0)\} \subset \mathbb{R}^2$ but eliminates examples such as $M = (\{0\}\times(-1,1)) \cup \{(x,\sin 1/x : x > 0)\}$. Indeed the above definition ensures the local representation ψ is a homeomorphism of V onto the image $M \cap W$ and hence ψ is an open map onto its image; that is

4.2 $V_0 \subset V$, V_0 open $\Rightarrow \exists$ an open $W_0 \subset W$ with $\psi(V_0) = M \cap W_0$.

The tangent space T_yM of M at y is the subspace of $\mathbb{R}^{n+\ell}$ consisting of those $\tau \in \mathbb{R}^{n+\ell}$ such that $\tau = \dot\gamma(0)$ for some C^1 curve $\gamma : (-1,1) \to \mathbb{R}^{n+\ell}$, $\gamma(-1,1) \subset M$, $\gamma(0) = y$. One readily checks that

4.3 T_yM is a linear subspace of $\mathbb{R}^{n+\ell}$ with basis $D_1\psi(x), \ldots D_n\psi(x)$,

where ψ is any local representation as in 4.1 above with $\psi(x) = y$.

A function $f : M \to \mathbb{R}^{n+\ell}$ $(k \ge 0)$ is said to be C^ℓ $(\ell \le r)$ on M if f is the restriction to M of a C^ℓ function $\overline{f} : U \to \mathbb{R}^{n+\ell}$, where U is an open set in $\mathbb{R}^{n+\ell}$ such that $M \subset U$.

4.4 Remark (Local graphical representations for M): If $\psi : V \to W$ is the local representation for a C^r submanifold M as above, and if $y_0 \in M \cap W$ and $x_0 = \psi^{-1}(y_0) \in V$, then, since rank $D\psi(x_0) = n$, there must be indices $1 \le k_1 < k_2 < \cdots < k_n \le n+\ell$ with $\det(D_{k_i}\psi_j(x_0)) \neq 0$. Hence supposing for the moment that $k_1 = 1, k_2 = 2, \ldots, k_n = n$ and letting $\widetilde{\psi} = (\psi_1, \ldots, \psi_n)$, the inverse function theorem implies that there are open $V_0, U \subset \mathbb{R}^n$ with $x_0 \in V_0 \subset V$ such that $\widetilde{\psi}|V_0$ is a C^r diffeomorphism of V_0 onto U. Observe that then $G = \psi \circ (\widetilde{\psi}|V_0)^{-1} : U \to W$ has the form

$$G(x) = (x, u(x)),\ x \in U,$$

where $u : U \to \mathbb{R}^\ell$ is defined by $u = (\psi_{n+1}, \ldots, \psi_{n+\ell}) \circ (\widetilde{\psi}|V_0)^{-1}$. That is G is the "graph map" $x \mapsto (x, u(x))$ corresponding to the C^r functionn u, and by construction

$$G(U) = \operatorname{graph} u = \psi(V_0) = M \cap W_0,$$

where $y_0 \in W_0 \subset \mathbb{R}^{n+\ell}$ and W_0 is open by 4.2. Without the assumption $k_i = i$, $i = 1, \ldots, n$, this of course remains true modulo composition with a permutation map (permuting the coordinates $x_1, \ldots, x_{n+\ell}$ in $\mathbb{R}^{n+\ell}$ so that the coordinates $x_{k_1}, \ldots, x_{k_n}$ are moved to the first n slots) so for each $y_0 \in M$ there is an open W_0 with $y_0 \in W_0$ and

$$(\ddagger) \qquad M \cap W_0 = Q(\operatorname{graph} u)$$

for some orthogonal transformation Q (where Q is in fact just a permutation of coordinates in $\mathbb{R}^{n+\ell}$) and for some C^r vector function $u = (u_1, \ldots, u_\ell)$ defined on an open set $U \subset \mathbb{R}^n$. Thus M is a C^r submanifold of $\mathbb{R}^{n+\ell}$ if and only if M is locally representable, near each of its points, as the graph of a C^r function u; i.e. each $y_0 \in M$ lies on some open W_0 such that ($\ddagger$) holds, with $u = (u_1, \ldots, u_\ell)$ a C^r vector function on some open $U \subset \mathbb{R}^n$.

We next want to discuss some differentiability properties for locally Lipschitz maps $f : M \to \mathbb{R}^P$ with $P \ge 1$ and also the area formula in case $P \ge n$. Thus $f : M \to \mathbb{R}^P$ and for each $x \in M$ we assume there are $\rho, L > 0$ with

$$\textbf{4.5} \qquad |f(y) - f(z)| \le L|y - z| \quad y, z \in M \cap B_\rho(x).$$

First we discuss directional derivatives of such an f: For given $\tau \in T_yM$ the directional derivative $D_\tau f \in \mathbb{R}^P$ is defined by

$$\textbf{4.6} \qquad D_\tau f = \tfrac{d}{dt} f(\gamma(t))\big|_{t=0}$$

for any C^1 curve $\gamma : (-1,1) \to M$ with $\gamma(0) = y$, $\dot\gamma(0) = \tau$, whenever this derivative exists. Of course it is easy to see that existence and the actual value is independent of the particular curve γ we use to represent τ because if $\widetilde{\gamma}$ is another such curve then, by 4.5,

$$\textbf{4.7} \qquad \lim_{t \downarrow 0} t^{-1}|f(\gamma(t)) - f(\widetilde{\gamma}(t))| \le L \lim_{t\downarrow 0} t^{-1}|\gamma(t) - \widetilde{\gamma}(t)| = 0$$

because $\gamma(0) = \widetilde{\gamma}(0) (= y)$ and $\gamma'(0) = \widetilde{\gamma}'(0) (= \tau)$.

We claim that in fact there is a set E of $\mathcal{H}^n$-measure zero such that $\forall x \in M \setminus E$

$$\textbf{4.8} \qquad D_\tau f(x) \text{ exists and the map } \tau \mapsto D_\tau f(x) \text{ is a linear map } T_xM \to \mathbb{R}^P.$$

Indeed this follows directly from the Rademacher theorem in $\mathbb{R}^n$ proved in 1.4, as follows: Let $y \in M$ and let $\varphi : U \cap M \to V$ be a local coordinate transformation; thus the inverse $\psi : V \to \mathbb{R}^{n+\ell}$ is C^1 with $\psi(V) = M \cap U$. Then according to 1.4 there is $E_0 \subset V$ with $\mathcal{H}^n(E_0) = 0$ such that $f \circ \psi$ is differentiable at every point of $V \setminus E_0$ and in particular for every $\eta \in \mathbb{R}^n$ and $x \in V \setminus E_0$ we have $D_\eta(f \circ \psi)(x) = \frac{d}{dt} f(\psi(x + t\eta))|_{t=0}$ exists and is linear in η. But $\gamma(t) = \psi(x + t\eta)$ is a curve as in 4.6 with $\tau = \sum_{j=1}^n \eta_j D_j\psi(x)$, so in fact this says that the directional derivatives

$$\textbf{4.9} \qquad D_{\sum_{j=1}^n \eta_j D_j\psi(x)} f(y) \text{ exist and } = D_\eta(f \circ \psi)(x)$$

and are linear in η for all $x \in V \setminus E_0$ and $y = \psi(x) \in U \cap M \setminus \psi(E_0)$. Hence, since $D_1\psi(x), \ldots, D_n\psi_n(x)$ is a basis for $T_{\psi(x)}M$, this just says that indeed 4.8 does hold at points of $U \cap M \setminus \psi(E_0)$, and of course $\psi(E_0)$ is a set of $\mathcal{H}^n$-measure zero because ψ is locally Lipschitz on V.

Notice also that if in fact f is the restriction of a locally Lipschitz function $\overline{f}$ defined in an open set $W \supset M$ then (by the same argument as in 4.7 with $\widetilde{\gamma}(t) = x + t\tau$) we have

$$\textbf{4.10} \qquad \forall \tau \in T_yM : \quad D_\tau f(y) \text{ exists } \iff \tfrac{d}{dt}\overline{f}(x + t\tau)|_{t=0} \text{ exists,}$$

and in that case the two quantities are equal.

Taking the particular choice $\eta = e_i$, $y = \psi(x) \in U \cap M \setminus \psi(E_0)$ in 4.9, and letting $\tau_1, \ldots, \tau_n$ be an orthonormal basis for T_yM, so that then

$$D_i\psi(x) = \textstyle\sum_{\ell=1}^n D_i\psi(x) \cdot \tau_\ell \tau_\ell,$$

we have

$$D_i(f \circ \psi)(x) = \textstyle\sum_{k=1}^n D_{\tau_k} f(y) D_i\psi(x) \cdot \tau_\ell,$$

whence

$$D_i(f\circ\psi)(x) \cdot D_j(f\circ\psi)(x) = \textstyle\sum_{k,m=1}^n (D_i\psi(x)\cdot\tau_k)(D_j\psi(x)\cdot\tau_m) D_{\tau_k} f(y) \cdot D_{\tau_m} f(y)$$

Since $\det AB = \det A \det B$ for square matrices A, B, and

$$\textstyle\sum_{k=1}^{n} D_i\psi(x)\cdot\tau_k\, D_j\psi(x)\cdot\tau_k = D_i\psi(x)\cdot D_j\psi(x)$$

because $D_i\psi(x) \in \operatorname{span}\{\tau_1,\ldots,\tau_n\}$, this implies

4.11 $$J_{f\circ\psi}(x) = J_\psi(x)\, J_f(y),$$

where we define

4.12 $$J_f(y) = \sqrt{\det \mathcal{J}(x)},$$

where $\mathcal{J}(y) =$ the $n\times n$ matrix with $(D_{\tau_k} f(y)\cdot D_{\tau_m} f(y))$ in the k-th row and m-th column. Now in case $P \ge n$ we conclude that the area formula holds for f; thus if f is 1:1 and $A \subset M$ is $\mathcal{H}^n$-measurable then

$$\mathcal{H}^n(f(A)) = \int_A J_f\, d\mathcal{H}^n,$$

and more generally, for any non-negative $\mathcal{H}^n$-measurable function $h : M \to [0,\infty)$ we have

4.13 $$\int_{f(M)} h\circ f^{-1}\, d\mathcal{H}^n = \int_M h\, J_f\, d\mathcal{H}^n.$$

If f is not assumed to be 1:1 then we still have

4.14 $$\int_{f(A)} \mathcal{N}(A,f)\, d\mathcal{H}^n = \int_A J_f\, d\mathcal{H}^n,$$

where $\mathcal{N}(A,f) = \mathcal{H}^0(f^{-1}\{y\}\cap A)$.

We can of course now (for any $P \ge 1$) define the induced linear map $df_y^M : T_yM \to \mathbb{R}^P$ just as it is in $\mathbb{R}^n$ by

4.15 $$df_y^M(\tau) = D_\tau f(y), \quad \tau \in T_yM.$$

In case f is real-valued (i.e. $P = 1$) then we define the gradient $\nabla^M f$ of f by

4.16 $$\nabla^M f(y) = \textstyle\sum_{j=1}^{n}\big(D_{\tau_j} f(y)\big)\tau_j,\ y \in M,$$

where $\tau_1,\ldots,\tau_n$ is any orthonormal basis for T_yM. If we let $\nabla_j^M f \equiv e_j\cdot\nabla^M f$ ($e_j = j$-th standard basis vector in $\mathbb{R}^{n+\ell}$, $j = 1,\ldots,n+\ell$) then

4.17 $$\nabla f(y) = \textstyle\sum_{j=1}^{n+\ell}\nabla_j^M f(y) e_j.$$

If f is the restriction to M of a $C^1(U)$ function $\overline{f}$, where U is an open subset of $\mathbb{R}^{n+\ell}$ containing M, then

$$\nabla^M f(y) = \big(\nabla_{\mathbb{R}^{n+\ell}}\overline{f}(y)\big)^\top,\ y \in M,$$

where $\nabla_{\mathbb{R}^{n+\ell}}\overline{f}$ is the usual $\mathbb{R}^{n+\ell}$ gradient $\big(D_1\overline{f},\ldots,D_{n+\ell}\overline{f}\big)$ on U, and where $(\ )^\top$ means orthogonal projection of $\mathbb{R}^{n+\ell}$ onto T_yM.

Now given a vector function ("vector field") $X = (X^1,\ldots,X^{n+\ell}) : M \to \mathbb{R}^{n+\ell}$ with $X^j \in C^1(M)$, $j = 1,\ldots,n+\ell$, we define

4.18 $$\operatorname{div}_M X = \textstyle\sum_{j=1}^{n+\ell}\nabla_j^M X^j$$

on M. (Notice that we do *not* require $X_y \in T_yM$.) Then, at $y \in M$, we have

4.19 $$\begin{aligned}\operatorname{div}_M X &= \textstyle\sum_{j=1}^{n+\ell} e_j\cdot(\nabla^M X^j)\\ &= \textstyle\sum_{j=1}^{n+\ell} e_j\cdot\big(\sum_{i=1}^{n}(D_{\tau_i}X^j)\tau_i\big),\end{aligned}$$

so that (since $X = \sum_{j=1}^{n+\ell} X^j e_j$)

4.20 $$\operatorname{div}_M X = \textstyle\sum_{i=1}^{n}\tau_i\cdot D_{\tau_i}X,$$

where $\tau_1,\ldots,\tau_n$ is any orthonormal basis for T_yM.

The divergence theorem states that if the closure $\overline{M}$ of M is a smooth compact manifold with boundary $\partial M = \overline{M}\setminus M$, and if $X_y \in T_yM\ \forall y \in M$, then

4.21 $$\int_M \operatorname{div}_M X\, d\mathcal{H}^n = -\int_{\partial M} X\cdot\eta\, d\mathcal{H}^{n-1}$$

where η is the inward pointing unit co-normal of ∂M; that is, $|\eta| = 1$, η is normal to ∂M, tangent to M, and points into M at each point of ∂M.

4.22 Remarks: (1) M need *not* be orientable here.

(2) In general the closure $\overline{M}$ of M will not be a nice manifold with boundary; indeed it can certainly happen that $\mathcal{H}^n(\overline{M}\setminus M) > 0$. (For example consider the case when $M = \{(x,y) \in \mathbb{R}^2 : x > 0, y = \sin(1/x)\}$. M is a C^∞ 1-dimensional submanifold of $\mathbb{R}^2$ in the sense of the above definitions, but $\overline{M}\setminus M$ is the interval $\{0\}\times[-1,1]$ on the y-coordinate axis.) Nevertheless in the general case we still have (in place of 4.21)

(‡) $$\int_M \operatorname{div}_M X = 0$$

provided support $X\cap M$ *is a compact subset of* M and $X_y \in T_yM\ \forall y \in M$.

In case M is at least C^2 we define the second fundamental form of M at y to be the bilinear form

4.23 $$B_y : T_yM\times T_yM \to (T_yM)^\perp$$

such that

4.24 $$B_y(\tau,\eta) = -\textstyle\sum_{\alpha=1}^k (\eta\cdot D_\tau\nu^\alpha)\nu^\alpha\big|_y,\ \ \tau,\eta\in T_yM,$$

where $\nu^1,\ldots,\nu^k$ are (locally defined, near y) vector fields with $\nu^\alpha(z)\cdot\nu^\beta(z)=\delta_{\alpha\beta}$ and $\nu^\alpha(z)\in(T_zM)^\perp$ for every z in some neighborhood of y. Notice that if $\eta_1,\ldots,\eta_\ell$ are vectors in $\mathbb{R}^{n+\ell}$ such that $D_1\psi(x),\ldots,D_n\psi_n(x),\eta_1,\ldots,\eta_\ell$ are linearly independent, we can obtain such locally defined orthonormal C^1 normal vector fields $\nu^1,\ldots,\nu^\ell$ by the Gram-Schmidt orthogonalization process applied to the basis $D_1\psi(\xi),\ldots,D_n\psi_n(\xi),\eta_1,\ldots,\eta_\ell$ of $\mathbb{R}^{n+\ell}$ for ξ in a suitable neighborhood of x.

The geometric significance of B is as follows: If $\tau\in T_yM$ with $|\tau|=1$ and $\gamma:(-1,1)\to\mathbb{R}^{n+\ell}$ is a C^2 curve with $\gamma(0)=y$, $\gamma(-1,1)\subset M$, and $\dot\gamma(0)=\tau$, then

4.25 $$B_y(\tau,\tau)=(\ddot\gamma(0))^\perp,$$

which is just the normal component (relative to M) of the curvature of γ at 0, γ being considered as an ordinary space-curve in $\mathbb{R}^{n+\ell}$. (Thus $B_y(\tau,\tau)$ measures the "normal curvature" of M in the direction τ.) To check this, simply note that $\nu^\alpha(\gamma(t))\cdot\dot\gamma(t)\equiv 0$, $|t|<1$, because $\dot\gamma(t)\in T_{\gamma(t)}M$ and $\nu^\alpha(\gamma(t))\in(T_{\gamma(t)}M)^\perp$. Differentiating this relation with respect to t, we get (after setting $t=0$)

$$\nu^\alpha(y)\cdot\ddot\gamma(0)=-(D_\tau\nu^\alpha)\cdot\tau$$

and hence (multiplying by $\nu^\alpha(y)$ and summing over α) we have

$$\begin{aligned}(\ddot\gamma(0))^\perp &= -\textstyle\sum_{\alpha=1}^k(\tau\cdot D_\tau\nu^\alpha)\nu^\alpha(y)\\ &= B_y(\tau,\tau)\end{aligned}$$

as required. (Note that the parameter t here need *not* be arc-length for γ; it suffices that $\dot\gamma(0)=\tau$, $|\tau|=1$.) More generally, by a similar argument, if $\tau,\eta\in T_yM$ and if $\varphi:U\to\mathbb{R}^{n+\ell}$ is a C^2 mapping of a neighborhood U of 0 in $\mathbb{R}^2$ such that $\varphi(U)\subset M$, $\varphi(0)=y$, $\frac{\partial\varphi}{\partial x^1}(0,0)=\tau$, $\frac{\partial\varphi}{\partial x^2}(0,0)=\eta$, then

4.26 $$B_y(\tau,\eta)=-\Big(\frac{\partial^2\varphi}{\partial x^1\partial x^2}(0,0)\Big)^\perp.$$

Of course such maps φ do exist for any given $\tau,\eta\in T_yM$, so 4.26 implies in particular that $B_y(\tau,\eta)=B_y(\eta,\tau)$; that is B_y is a *symmetric* bilinear form with values in $(T_yM)^\perp$.

We define the mean curvature vector $\underline{H}$ of M at y to be trace B_y; thus

4.27 $$\underline{H}(y)=\textstyle\sum_{i=1}^n B_y(\tau_i,\tau_i)\in(T_yM)^\perp,$$

where $\tau_1,\ldots,\tau_2$ is an orthonormal basis for T_yM. Notice that then (if $\nu^1,\ldots,\nu^k$ are as above)

$$\underline{H}(y)=-\textstyle\sum_{\alpha=1}^k\sum_{i=1}^n(\tau_i\cdot D_{\tau_i}\nu^\alpha)\nu^\alpha(y)$$

so that

4.28 $$\underline{H}(y)=-\textstyle\sum_{\alpha=1}^k(\operatorname{div}_M\nu^\alpha)\nu^\alpha$$

near y.

Returning for a moment to 4.21 (in case $\overline{M}$ is a compact C^2 manifold with smooth $(n-1)$-dimensional boundary $\partial M=\overline{M}\setminus M$) it is interesting to compute $\int_M\operatorname{div}_M X$ in case the condition $X_y\in T_yM$ is *dropped*. To compute this, we decompose X into its tangent and normal parts:

$$X=X^\top+X^\perp$$

where (at least locally, in the notation introduced of 4.24 above)

$$X^\perp=\textstyle\sum_{\alpha=1}^k(\nu^\alpha\cdot X)\nu^\alpha.$$

Then we have (near y)

$$\operatorname{div}_M X^\perp=\textstyle\sum_{\alpha=1}^k(\nu^\alpha\cdot X)\operatorname{div}\nu^\alpha,$$

so that by 4.28

4.29 $$\operatorname{div}_M X^\perp=-X\cdot\underline{H}$$

at each point of M. On the other hand $\int_M\operatorname{div}_M X^\top=-\int_{\partial M}X\cdot\eta$ by 4.21. Hence, since $\operatorname{div}_M X=\operatorname{div}_M X^\top+\operatorname{div}_M X^\perp$, we obtain

4.30 $$\int_M\operatorname{div}_M X\,d\mathcal{H}^n=-\int_M X\cdot\underline{H}\,d\mathcal{H}^n-\int_{\partial M}X\cdot\eta\,d\mathcal{H}^{n-1}.$$

5 First and Second Variation Formulae

Suppose that M is an n-dimensional C^1 submanifold of $\mathbb{R}^{n+\ell}$ and let U be an open subset of $\mathbb{R}^{n+\ell}$ such that $U\cap M\neq\emptyset$ and $\mathcal{H}^n(K\cap M)<\infty$ for each compact $K\subset U$. Also, let $\{\varphi_t\}_{-1\le t\le 1}$ be a 1-parameter family of diffeomorphisms $U\to U$ such that, for some compact subset $K\subset U$,

5.1 $$\begin{cases}\varphi_t(x))=\varphi(t,x)\text{ is a }C^2\text{ map of }(t,x)\in(-1,1)\times U\to U\\ \varphi_0(x)=x,\quad\forall x\in U\\ \varphi_t(x)=x,\quad\forall(t,x)\in(-1,1)\times U\setminus K,\end{cases}$$

Also, let X, Z denote the initial velocity and acceleration vectors for φ_t: thus $X_x = \frac{\partial\varphi(t,x)}{\partial t}\Big|_{t=0}$, $Z_x = \frac{\partial^2\varphi(t,x)}{\partial t^2}\Big|_{t=0}$. Then

5.2
$$\varphi_t(x) = x + tX_x + \frac{t^2}{2}Z_x + O(t^3)$$

and X, Z have supports which are compact subsets of U. Let $M_t = \varphi_t(M \cap K)$ (K as in 5.1); thus M_t is a 1-parameter family of manifolds such that $M_0 = M \cap K$ and M_t agrees with M outside some compact subset of U. We want to compute $\frac{d}{dt}\mathcal{H}^n(M_t)\Big|_{t=0}$ and $\frac{d^2}{dt^2}\mathcal{H}^n(M_t)\Big|_{t=0}$ (i.e. the "first and second variation" of M). The area formula is particularly useful here because it gives (with K as in 5.1)

$$\mathcal{H}^n(\varphi_t(M \cap K)) = \int_{M\cap K} J_{\psi_t}\, d\mathcal{H}^n,\ \psi_t = \varphi_t|M \cap U,$$

and hence to compute the first and second variation we can differentiate under the integral. Thus the computation reduces to calculation of $\frac{\partial}{\partial t}J_{\psi_t}\Big|_{t=0}$ and $\frac{\partial^2}{\partial t^2}J_{\psi_t}\Big|_{t=0}$. To make the calculation we need to get an explicit expression for the terms up to second order in the Taylor series expansion (in the variable t) of J_{ψ_t}. Note that (for fixed x)

$$\begin{aligned} d\psi_{t|x}(\tau) &= D_\tau\psi_t \quad (\tau \in T_xM)\\ &= \tau + tD_\tau X + \frac{t^2}{2}D_\tau Z + O(t^3) \quad \text{by 5.2.}\end{aligned}$$

Hence, relative to the orthonormal bases $\tau_1,\ldots,\tau_n$ for T_xM and $e_1,\ldots,e_{n+\ell}$ for $\mathbb{R}^{n+\ell}$, the map $d\psi_{t|x} : T_xM \to \mathbb{R}^{n+\ell}$, has matrix with i-th row

$$D_{\tau_i}\psi_t(x) = \tau_i + tD_{\tau_i}X + \frac{t^2}{2}D_{\tau_i}Z + O(t^3)$$

for $i = 1,\ldots,n$. Then, with respect to $\tau_1,\ldots,\tau_n$, $(d\psi_{t|x})^*\circ(d\psi_{t|x})$ has matrix (b_{ij}),

$$\begin{aligned} b_{ij} &= D_{\tau_i}\psi_{t|x}\cdot D_{\tau_j}\psi_{t|x}\\ &= \delta_{ij} + t(\tau_i\cdot D_{\tau_j}X + \tau_j\cdot D_{\tau_i}X)\\ &\quad + t^2\big(\tfrac{1}{2}(\tau_i\cdot D_{\tau_j}Z + \tau_j\cdot D_{\tau_i}Z) + (D_{\tau_i}X)\cdot(D_{\tau_j}X)\big) + O(t^3),\end{aligned}$$

so that, by the general formula

$$\det(I + A) = 1 + \operatorname{trace} A + \tfrac{1}{2}(\operatorname{trace} A)^2 - \tfrac{1}{2}\operatorname{trace}(A^2) + O(|A|^3),$$

$$\begin{aligned}(J_{\psi_t})^2 &= 1 + 2t\operatorname{div}_M X + t^2\big(\operatorname{div}_M Z + \textstyle\sum_{i=1}^n|D_{\tau_i}X|^2\\ &\qquad + 2(\operatorname{div}_M X)^2 - \tfrac{1}{2}\textstyle\sum_{i,j=1}^n(\tau_i\cdot D_{\tau_j}X + \tau_j\cdot D_{\tau_i}X)^2\big) + O(t^3)\\ &= 1 + 2t\operatorname{div}_M X + t^2\big(\operatorname{div}_M Z + \textstyle\sum_{i=1}^n|(D_{\tau_i}X)^\perp|^2\\ &\qquad + 2(\operatorname{div}_M X)^2 - \textstyle\sum_{i,j=1}^n(\tau_i\cdot D_{\tau_j}X)(\tau_j\cdot D_{\tau_i}X)\big) + O(t^3),\end{aligned}$$

where $(D_{\tau_i}X)^\perp$ (= the normal part of $D_{\tau_i}X$) $= D_{\tau_i}X - \sum_{j=1}^n(\tau_j\cdot D_{\tau_i}X)\tau_j$. Using $\sqrt{1+x} = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + O(x^3)$, we thus get

$$\begin{aligned} J_{\psi_t} &= 1 + t\operatorname{div}_M X + \tfrac{t^2}{2}\big(\operatorname{div}_M Z + (\operatorname{div}_M X)^2 + \textstyle\sum_{i=1}^n|(D_{\tau_i}X)^\perp|^2\\ &\qquad - \textstyle\sum_{i,j=1}^n(\tau_i\cdot D_{\tau_j}X)(\tau_j\cdot D_{\tau_i}X)\big) + O(t^3).\end{aligned}$$

Thus the area formula immediately yields the first variation formula

5.3
$$\frac{d}{dt}\mathcal{H}^n(M_t)\Big|_{t=0} = \int_M \operatorname{div}_M X\, d\mathcal{H}^n$$

and the second variation formula

5.4
$$\begin{aligned}\frac{d^2}{dt^2}\mathcal{H}^n(M_t)\,\Big|_{t=0} &= \int_M\big(\operatorname{div}_M Z + (\operatorname{div}_M X)^2 + \textstyle\sum_{i=1}^n|(D_{\tau_i}X)^\perp|^2\\ &\qquad - \textstyle\sum_{i,j=1}^n(\tau_i\cdot D_{\tau_j}X)(\tau_j\cdot D_{\tau_i}X)\big)\,d\mathcal{H}^n.\end{aligned}$$

We shall use the terminology that M *is stationary in* U if $\mathcal{H}^n(M \cap K) < \infty$ for each compact $K \subset U$ and if $\frac{d}{dt}\mathcal{H}^n(M_t)\Big|_{t=0} = 0$ where $M_t = \varphi_t(M \cap K)$, $\varphi_t(x) \equiv x$ for all $(x,t) \in (U \setminus K)\times(-\varepsilon,\varepsilon)$. Notice that by the first variation formula 5.3 this is equivalent to $\int_M \operatorname{div}_M X\, d\mathcal{H}^n = 0$ whenever X is C^1 on U with support X a compact subset of U.

In view of 4.30 we also have the following:

5.5 Lemma. *Suppose M is a C^2 submanifold of $\mathbb{R}^{n+\ell}$ with mean curvature vector $\underline{H}$ and $U \subset \mathbb{R}^{n+\ell}$ is open. Then*

(1) *If $\overline{M}$ is a C^2 submanifold with smooth $(n-1)$-dimensional boundary $\partial M = \overline{M}\setminus M$, then M is stationary in U if and only if $\underline{H} \equiv 0$ on $M \cap U$ and $\partial M \cap U = \emptyset$.*

(2) *If $\overline{U}\cap M$ is a compact subset of M, then M is stationary in U if and only if $\underline{H} \equiv 0$ on $M \cap U$.*

For later reference we also want to mention an important modification of the idea that M is stationary in U with U is open in $\mathbb{R}^{n+\ell}$. Namely, suppose N is a C^2 $(n+\ell)$-dimensional submanifold of $\mathbb{R}^{n+L}$, $0 \le \ell \le L$, and suppose U is an open subset of $\mathbb{R}^{n+L}$ such that $(\overline{N}\setminus N)\cap U = \emptyset$, and let $\{\varphi_t\}_{-1\le t\le 1}$ a 1-parameter family of diffeomorphisms $U \to U$ such that

5.6
$$\begin{cases} \varphi_t(x)) = \varphi(t,x) \text{ is a } C^2 \text{ map of } (t,x) \in (-1,1)\times U \to U\\ \qquad\qquad \text{with } \varphi_t(U\cap N) \subset U \cap N\ \forall t \in (-1,1)\\ \varphi_0(x) = x\ \forall x \in U, \quad \varphi_t(x) = x\ \forall (t,x) \in (-1,1)\times U\setminus K,\end{cases}$$

where K is a compact subset of U. Then we have the following definition:

5.7 Definition: V is stationary in $U \cap N$ if $\frac{d}{dt}\mathbb{M}(\phi_{t\#}(V \llcorner K))\Big|_{t=0} = 0$ whenever the family $\{\phi_t\}_{t\in(-1,1)}$ is as in 5.6.

In view of the fact that if X is any C^1 vector field in U with compact support K such that $K \cap N$ is a compact subset of N and such that $X_x \in T_x N$ at each $x \in N$, then there is a 1-parameter family φ_t as in 5.6 above with $\frac{\partial}{\partial t}\varphi(x,t)|_{t=0} = X_x$ at each point $x \in N \cap U$, we see that M is stationary in $U \cap N$ as in 5.7 if and only if

5.8 $$\int_M \operatorname{div}_M X = 0$$

whenever X is a C^1 vector field in U with compact support K such that $K \cap N$ is compact and whenever $X_x \in T_x N\ \forall x \in M$.

If we let $\nu^1, \ldots, \nu^L$ be an orthonormal family (defined locally near a point $y \in M$) of vector fields normal to M, such that $\nu^1, \ldots, \nu^\ell$ are tangent to N and $\nu^{\ell+1}, \ldots, \nu^L$ are normal to N, then for any vector field X on M we can write $X = X^T + X^\perp$, where $X_z^T \in T_z N$ and $X^\perp = \sum_{j=\ell+1}^k (\nu^j \cdot X)\nu^j$ (= the part of X normal to N). Then if $\tau_1, \ldots, \tau_n$ is any orthonormal basis for $T_y M$, we have

5.9 $$\begin{aligned}\operatorname{div}_M X &= \operatorname{div}_M X^T + \textstyle\sum_{j=\ell+1}^L (\nu^j \cdot X)\operatorname{div}_M \nu^j \\ &= \operatorname{div}_M X^T + \textstyle\sum_{j=\ell+1}^L (\nu^j \cdot X)\sum_{i=1}^n \tau_i \cdot D_{\tau_i}\nu^j \\ &= \operatorname{div}_M X^T - \textstyle\sum_{i=1}^n X \cdot \overline{B}_y(\tau_i, \tau_i),\end{aligned}$$

where $\overline{B}_y$ is the second fundamental form of N at y and where we used the definition of second fundamental form as in 4.24 (with N in place of M) and hence by virtue of 5.8 (with $X^{(1)}$ in place of X) we conclude:

5.10 Lemma. *If N is an $(n+\ell)$-dimensional C^2 submanifold of $\mathbb{R}^{n+L}$, if $M \subset N$ and if U is an open subset of $\mathbb{R}^{n+L}$ such that $\mathcal{H}^n(M \cap K) < \infty$ whenever K is a compact subset of U, then M is stationary in $U \cap N$ (in the sense of* Definition 5.7*) if and only if*

$$\int_M \operatorname{div}_M X = -\int_M \overline{H}_M \cdot X$$

for each C^1 vector field X with compact support contained in U; here

$$\overline{H}_M|_y = \textstyle\sum_{i=1}^n \overline{B}_y(\tau_i, \tau_i), \quad y \in M,$$

where $\overline{B}_y$ denotes the second fundamental form of N at y and $\tau_1, \ldots, \tau_n$ is any orthonormal basis of $T_y M$.

Finally, we shall need later the following important fact about the second variation formula 5.4.

5.11 Lemma. *If M is C^2, stationary in U, U open in $\mathbb{R}^{n+\ell}$ with $(\overline{M} \setminus M) \cap U = \emptyset$, and if X as in* 5.4 *has compact support in U with $X_y \in (T_y M)^\perp\ \forall y \in M$, then* 5.4 *says*

$$\frac{d^2}{dt^2}\mathcal{H}^n(M_t)\Big|_{t=0} = \int_M \big(\textstyle\sum_{i=1}^n |(D_{\tau_i}X)^\perp|^2 - \sum_{i,j=1}^n (X \cdot B(\tau_i, \tau_j))^2\big)\, d\mathcal{H}^n.$$

5.12 Remark: In case $L = 1$ and M is orientable, with continuous unit normal ν, then $X = \zeta\nu$ for some scalar function ζ with compact support on M, and the above identity has the simple form

$$(*)\qquad \frac{d^2}{dt^2}\mathcal{H}^n(M_t)\Big|_{t=0} = \int_M (|\nabla^M \zeta|^2 - \zeta^2|B|^2)\, d\mathcal{H}^n,$$

where $|B|^2 = \sum_{i,j=1}^n |B(\tau_i, \tau_j)|^2 \equiv \sum_{i,j=1}^n |\nu \cdot B(\tau_i, \tau_j)|^2$. This is clear, because $(D_{\tau_i}(\nu\zeta))^\perp = \nu D_{\tau_i}\zeta$ by virtue of the fact that $D_{\tau_i}\nu\big|_y \in T_y M\ \forall y \in M$.

Proof of 5.11: First we note that $\int_M \operatorname{div}_M Z\, d\mathcal{H}^n = 0$ by virtue of the fact that M is stationary in U, and $\operatorname{div}_M X = -X \cdot \underline{H} = 0$ by virtue of 4.29 and 5.5(2) and the fact that X s normal to M. The proof is then completed by noting that $\tau_i \cdot D_{\tau_j} X = -X \cdot B(\tau_i, \tau_j)$ by virtue of 4.24 and the fact that X is normal to M. □

6 Co-Area Formula and C^1 Sard Theorem

As in our discussion of the area formula, we begin by looking at linear maps $\lambda : \mathbb{R}^n \to \mathbb{R}^m$, but here we assume $m < n$, so $n = m + k$ with $k \in \{1, 2, \ldots\}$. Let us first look at the special case when λ is the orthogonal projection p of $\mathbb{R}^n$ onto $\mathbb{R}^m$. The orthogonal projection $p : \mathbb{R}^n \to \mathbb{R}^m$ (i.e. $p : x = (x^1, \ldots, x^n) \mapsto (x^1, \ldots, x^m)$) has the property that, $p^{-1}(0)$ is a k-dimensional subspace. Thus the inverse images $p^{-1}(y)$ are k-dimensional affine spaces, each being a translate of the k-dimensional subspace $p^{-1}(0)$. Thus the inverse images $p^{-1}(y)$ decompose all of $\mathbb{R}^n$ into parallel "k-dimensional slices" and by Fubini's Theorem

6.1 $$\int_{\mathbb{R}^m} \mathcal{H}^k(p^{-1}(y) \cap A)\, d\mathcal{L}^m(y) = \mathcal{H}^n(A)$$

whenever A is an $\mathcal{L}^n$-measurable subset of $\mathbb{R}^n$.

This formula (which, we emphasize again, is just Fubini's Theorem) is a special case of a more general formula known as the co-area formula. We first derive this in case of an arbitrary linear map $\lambda : \mathbb{R}^n \to \mathbb{R}^m$ with rank $\lambda = m$.

Let $F = \lambda^{-1}(0)$. (Then for each $y \in \mathbb{R}^m$, $\lambda^{-1}(y)$ is a k-dimensional affine space which is a translate of F; the sets $\lambda^{-1}(y)$ thus decompose all of $\mathbb{R}^n$ into parallel k-dimensional slices, $k = n - m$)

Take an orthogonal transformation $q = \mathbb{R}^n \to \mathbb{R}^n$ such that $qF^\perp = \mathbb{R}^m \times \{0\}$, $qF = \{0\} \times \mathbb{R}^k$. Then $\lambda \circ q^*$ is zero on $\{0\} \times \mathbb{R}^k$ and is an isomorphism of $\mathbb{R}^m \times \{0\}$ onto $\mathbb{R}^m$, hence $\lambda = \sigma \circ p \circ q$, where $p(x, y) = x$ is as above and σ is a linear isomorphism of $\mathbb{R}^m$. By 6.1 above, for any $\mathcal{H}^n$-measurable $A \subset \mathbb{R}^n$,

$$\begin{aligned}\mathcal{L}^n(A) = \mathcal{L}^n(q(A)) &= \int_{\mathbb{R}^m} \mathcal{H}^k\big(q(A) \cap p^{-1}(y)\big)\, d\mathcal{L}^m(y)\\ &= \int_{\mathbb{R}^m} \mathcal{H}^k\big(A \cap q^{-1}(p^{-1}(y))\big)\, d\mathcal{L}^m(y).\end{aligned}$$

Making the change of variable $z = \sigma(y)$ $dy = |\det\sigma|^{-1}\, dz$, we thus get

$$\begin{aligned}|\det\sigma|\mathcal{L}^n(A) &= \int_{\mathbb{R}^m} \mathcal{H}^k\big(A \cap q^{-1}(p^{-1}(\sigma^{-1}(z)))\big)\, d\mathcal{L}^m(z)\\ &= \int_{\mathbb{R}^m} \mathcal{H}^k\big(A \cap \lambda^{-1}(z)\big)\, d\mathcal{L}^m(z).\end{aligned}$$

Also, since $q^* \circ q = 1_{\mathbb{R}^n}$ and $p \circ p^*$ is the identity on $\mathbb{R}^m$, we have $\lambda \circ \lambda^* = \sigma \circ \sigma^* : \mathbb{R}^m \to \mathbb{R}^m$, so that $|\det\sigma| = \sqrt{\det \lambda \circ \lambda^*}$.

Thus finally

6.2 $$\sqrt{\det \lambda \circ \lambda^*}\, \mathcal{L}^n(A) = \int_{\mathbb{R}^m} \mathcal{H}^k\big(A \cap \lambda^{-1}(y)\big)\, d\mathcal{L}^m(y).$$

This is the co-area formula for linear maps. (Note that it is trivially valid, with both sides zero, in case rank $\lambda < m$.)

Generally, given a C^1 map $f : M \to \mathbb{R}^m$, where M is an n-dimensional C^1 submanifold of some Euclidean space $\mathbb{R}^N$, we can define

$$J_f^*(x) = \sqrt{\det(df_x) \circ (df_x)^*},$$

where, as usual, $df_x : T_x M \to \mathbb{R}^m$ denotes the induced linear map. Then for any Borel set $A \subset M$

6.3 $$\int_A J_f^*\, d\mathcal{H}^n = \int_{\mathbb{R}^m} \mathcal{H}^k\big(A \cap f^{-1}(y)\big)\, d\mathcal{L}^m(y).$$

This is the general co-area formula. Its proof uses an approximation argument based on the linear case 6.2. (See [Har79] or [Fed69] for the details.)

An important consequence of 6.3 is that if $C = \{x \in M : J_f^*(x) = 0\}$, then (by using 6.3 with $A = C$) $\mathcal{H}^k(C \cap f^{-1}(y)) = 0$ for $\mathcal{L}^m$-a.e. $y \in \mathbb{R}^m$. Also, since $J_f^*(x) \neq 0$ precisely when df_x has rank m, the implicit function theorem implies that either $f^{-1}(y) \setminus C$ is empty or else it is an k-dimensional C^1 submanifold in the sense of §4 above.

In summary we thus have the following important result.

6.4 Theorem (C^1 Sard-type Theorem.) *Suppose $f : M \to \mathbb{R}^m$, $m < n$, is C^1, with M is an n-dimensional C^1 submanifold of $\mathbb{R}^N$. Then for $\mathcal{L}^m$-a.e. $y \in f(M)$, $f^{-1}(y)$ decomposes into an k-dimensional C^1 submanifold and a closed set of $\mathcal{H}^k$-measure zero, where $k = n - m$. Specifically,*

$$f^{-1}(y) = \big(f^{-1}(y) \setminus C\big) \cup \big(f^{-1}(y) \cap C\big),$$

$C = \{x \in M : J_f^*(x) = 0\}$ ($\equiv \{x \in M : \operatorname{rank}(df_x) < m\}$), $\mathcal{H}^k\big(f^{-1}(y) \cap C\big) = 0$, $\mathcal{L}^m$-a.e. y, and $f^{-1}(y) \setminus C$ is either empty or an k-dimensional C^1 submanifold.

6.5 Remark: If f and M are of class C^{k+1}, then *Sard's Theorem* asserts the stronger result that in fact $f^{-1}(y) \cap C = \emptyset$ for $\mathcal{L}^m$-a.e. $y \in \mathbb{R}^m$, so that $f^{-1}(y)$ is a k-dimensional C^{k+1} submanifold for $\mathcal{L}^m$-a.e. $y \in \mathbb{R}^m$.

A useful generalization of 6.3, obtained by applying 6.3 to a suitable increasing sequence of simple functions, is as follows: If g is a non-negative $\mathcal{H}^n$-measurable function on M, then

6.6 $$\int_M g J_f^*\, d\mathcal{H}^n = \int_{\mathbb{R}^m} \int_{f^{-1}(y)} g\, d\mathcal{H}^k\, d\mathcal{L}^m(y).$$

6.7 Remarks: (1) Notice that the above formulae enable us to bound the $\mathcal{H}^k$ measure of the "slices" $f^{-1}y$ for a good set of y. Specifically if $|f| \le R$ and g is as in 6.6 ($g \equiv 1$ is an important case), then there must be set $S \subset B_R(0)$ ($\subset \mathbb{R}^m$), $S = S(g, f, M)$, with $\mathcal{L}^m(S) \ge \frac{1}{2}\mathcal{L}^m(B_R(0))$ and with

$$\int_{f^{-1}(y)} g\, d\mathcal{H}^k \le \frac{2}{\mathcal{L}^m(B_R(0))} \int_M g\, J_f^*\, d\mathcal{H}^n$$

for each $y \in S$. For otherwise there would be a set $T \subset B_R(0)$ with $\mathcal{L}^m(T) > \frac{1}{2}\mathcal{L}^m(B_R(0))$ and

$$\int_{f^{-1}(y)} g\, d\mathcal{H}^k \ge \frac{2}{\mathcal{L}^m(B_R(0))} \int_M g\, J_f^*\, d\mathcal{H}^n,\ y \in T,$$

so that, integrating over T we obtain a contradiction to 6.6 if $\int_M g\, J_f^*\, d\mathcal{H}^n > 0$. On the other hand if $\int_M q\, J_f^*\, d\mathcal{H}^n = 0$ then the required result is a trivial consequence of 6.6.

(2) The above has an important extension to the case when we have $f : \mathbb{R}^N \to \mathbb{R}^n$ and *sequences* $\{M_j\}$, $\{g_j\}$ satisfying the conditions of M, g above. In this case there is a set $S \subset B_R(0)$ with $\mathcal{L}^m(S) \ge \frac{1}{2}\mathcal{L}^m(B_R(0))$ such that for each $y \in S$ there is a *subsequence* $\{j'\}$ (depending on y) with

$$\int_{M_{j'}\cap f^{-1}(y)} g_{j'}\, d\mathcal{H}^k \le \frac{2}{\mathcal{L}^m(B_R(0))}\int_{M_{j'}} g_{j'}\, J_f^*\, d\mathcal{H}^n.$$

Indeed otherwise there is a set T with $\mathcal{L}^m(T) > \frac{1}{2}\mathcal{L}^m(B_R(0))$ so that for each $y \in T$ there is $\ell(y)$ such that

$$(*) \qquad \int_{M_j\cap f^{-1}(y)} g_j\, d\mathcal{H}^k > \frac{2}{\mathcal{L}^m(B_R(0))}\int_{M_j} g_j\, J_f^*\, d\mathcal{H}^n$$

for each $j > \ell(y)$. But $T = \cup_{j=1}^\infty T_j$, $T_j = \{y \in T : \ell(y) \le j\}$, and hence there must exist j so that $\mathcal{L}^m(T_j) > \frac{1}{2}\mathcal{L}^m(B_R(0))$. Then, integrating $(*)$ over $y \in T_j$, we obtain a contradiction to 6.6 as before.

Chapter 3

Countably *n*-Rectifiable Sets

1 Basic Notions, Tangent Properties

Firstly, a set $M \subset \mathbb{R}^{n+\ell}$ is said to be countably n-rectifiable if

1.1 $$M \subset M_0 \cup \left(\cup_{j=1}^\infty F_j(\mathbb{R}^n)\right),$$

where $\mathcal{H}^n(M_0) = 0$ and $F_j : \mathbb{R}^n \to \mathbb{R}^{n+\ell}$ are Lipschitz functions for $j = 1, 2, \ldots$.[1] Notice also that by the extension theorem 1.2 of Ch. 2 this is equivalent to saying

$$M = M_0 \cup \left(\cup_{j=1}^\infty F_j(A_j)\right)$$

where $\mathcal{H}^n(M_0) = 0$, $F_j : A_j \to \mathbb{R}^{n+\ell}$ Lipschitz, $A_j \subset \mathbb{R}^n$. More importantly, we have the following lemma.

1.2 Lemma. *M is countably n-rectifiable if and only if $M \subset \cup_{j=0}^\infty N_j$, where $\mathcal{H}^n(N_0) = 0$ and where each N_j, $j \ge 1$, is an n-dimensional embedded C^1 submanifold of $\mathbb{R}^{n+\ell}$.*

Proof: The "if" part is essentially trivial because if N is an n-dimensional C^1 submanifold, then using local representations for N as in Remark 4.4 of Ch. 2 we see

[1]Notice that this differs slightly from the terminology of [Fed69] in that we allow a set M_0 with $\mathcal{H}^n(M_0) = 0$.

that for each $x \in N$ there is $\rho_x > 0$ such that $B_{\rho_x}(x) \cap N = \psi(V)$ for suitable C^1 map $\psi : V \to \mathbb{R}^{n+\ell}$, $V \subset \mathbb{R}^n$ open. Since such C^1 maps are automatically Lipschitz in each closed ball $\subset V$ it is then clear that M satisfies the definition 1.1.

The "only if" part is a consequence of the C^1 Approximation Theorem 1.5 of Ch. 2, which says that for each $j = 1, 2, \ldots$ we can choose C^1 functions $G_{1j}, G_{2j}, \ldots : \mathbb{R}^n \to \mathbb{R}^\ell$ such that, if F_j are Lipschitz functions as in the Definition 1.1 above, then $\mathcal{H}^n(\{x : F_j(x) \neq G_{ij}(x)\}) < 1/i$. So, with

$$Z_j = \mathbb{R}^n \setminus (\cup_{i=1}^\infty \{x : F_j(x) = G_{ij}(x)\}),$$

we have $\mathcal{H}^n(Z_j) = 0$, in which case

$$\text{(1)} \qquad \text{graph } F_j \subset E_j \cap \left(\cup_{i=1}^\infty G_{ij}(\mathbb{R}^n)\right), \quad j = 1, 2, \ldots$$

where $E_j = F_j(Z_j)$. Then $\mathcal{H}^n(E_j) = 0$ because F_j is Lipschitz and $\mathcal{H}^n(Z_j) = 0$, so

$$\text{(2)} \qquad \mathcal{H}^n(N_0) = 0, \text{ where } N_0 = \left(\cup_{j=1}^\infty E_j\right),$$

and we have proved

$$M \subset M_0 \cup N_0 \cup (\cup_{i,j=1}^\infty G_{ij}\mathbb{R}^n).$$

Now by the area formula $\mathcal{L}^n(\{x : J_{G_{ij}} = 0\}$ whereas if the Jacobian G_{ij} is non-zero at a point x, then there is a $\rho > 0$ such that $G_{ij}(\breve{B}_\rho(x))$ is an n-dimensional C^1 submanifold of $\mathbb{R}^n$ (with G_{ij} providing a local representation in a neighborhood of the point $y = G_{ij}(x)$). So $\cup_{ij} G_{ij}$ can be written as the union of a set of measure zero and countably many n-dimensional C^1 submanifolds of $\mathbb{R}^{n+\ell}$. □

1.3 Remark: If M is countably n-rectifiable, the above lemma guarantees that we can find N_0 with $\mathcal{H}^n$ measure zero and n-dimensional C^1 submanifolds $N_1, N_2, \ldots$ with $M \subset \cup_{j=0}^\infty N_j$, and so we can write M as a disjoint union $M = \cup_{j=0}^\infty M_j$ with $M_j \subset N_j$ for each $j = 0, 1, 2, \ldots$. To achieve this, just define the M_j inductively by $M_0 = M \cap N_0$ and $M_j = M \cap N_j \setminus \cup_{i=0}^{j-1} M_i$, $j \geq 1$. Of course the sets M_j so constructed are all $\mathcal{H}^n$-measurable if M is.

We now want to give an important characterization of countably n-rectifiable sets in terms of *approximate tangent spaces*, which we first define:

1.4 Definition: If M is an $\mathcal{H}^n$-measurable subset of $\mathbb{R}^{n+\ell}$ with $\mathcal{H}(M \cap K) < \infty\ \forall$ compact K, then we say that an n-dimensional subspace of P of $\mathbb{R}^{n+\ell}$ is the approximate tangent space for M at x (x a given point in $\mathbb{R}^{n+\ell}$) if

$$\lim_{\lambda \downarrow 0} \int_{\eta_{x,\lambda}(M)} f(y)\, d\mathcal{H}^n(y) = \int_P f(y)\, d\mathcal{H}^n(y)\ \forall f \in C_c^0(\mathbb{R}^{n+\ell}).$$

(Recall $\eta_{x,\lambda} : \mathbb{R}^{n+\ell} \to \mathbb{R}^{n+\ell}$ is defined by $\eta_{x,\lambda}(y) = \lambda^{-1}(y - x)$, $x, y \in \mathbb{R}^{n+\ell}$, $\lambda > 0$.)

1.5 Remarks: (1) Of course P is unique if it exists; we shall denote it by $T_x M$.

(2) We shall show below (in the proof of the "⇒" part of 1.6) that, with M_j, N_j as in Remark 1.3 above,

$$T_x M = T_x N_j,\ \mathcal{H}^n\text{-a.e. } x \in M_j,\ j = 1, 2, \ldots.$$

This is a very useful fact.

(3) By choosing $f : \mathbb{R}^{n+\ell} \to [0, 1] \in C_c^0(\mathbb{R}^{n+\ell})$ with $f \equiv 1$ on $B_1(0)$ and $f \equiv 0$ on $\mathbb{R}^{n+\ell} \setminus B_{1+\varepsilon}(0)$ in Definition 1.4, we see (after letting $\varepsilon \downarrow 0$) that $T_x M$ exists ⇒

$$\lim_{\rho \downarrow 0} (\omega_n \rho^n)^{-1} \mathcal{H}^n(M \cap B_\rho(x)) = 1.$$

and similarly by taking functions $f : \mathbb{R}^{n+\ell} \to [0, 1] \in C_c^0(\mathbb{R}^{n+\ell})$ approximating the indicator function $\chi_{X_{1/2}((T_x M)^\perp, 0) \cap B_1(0)}$ we see that $T_x M$ exists ⇒

$$\lim_{\rho \downarrow 0} \frac{\mu\left(X_{\frac{1}{2}}((T_x M)^\perp, x) \cap B_\rho(x)\right)}{\omega_n \rho^n} = 0.$$

The following theorem gives the important characterization of countably n-rectifiable sets in terms of existence of approximate tangent spaces.

1.6 Theorem. *Suppose M is $\mathcal{H}^n$-measurable with $\mathcal{H}^n(M \cap K) < \infty$ for each compact $K \subset \mathbb{R}^{n+\ell}$. Then M countably n-rectifiable $\iff$ the approximate tangent space $T_x M$ exists for $\mathcal{H}^n$-a.e. $x \in M$.*

Proof of 1.6 "⇒": As described in Remark 1.3 above, we may write M as the disjoint union $\cup_{j=0}^\infty M_j$, where $\mathcal{H}^n(M_0) = 0$, $M_j \subset N_j$, $j \geq 1$, N_j embedded C^1 submanifolds of dimensions n, and M_j $\mathcal{H}^n$-measurable. Let $R > 0$ $f \in C_c^0(\mathbb{R}^{n+\ell})$ with $f \equiv 0$ in $\mathbb{R}^{n+\ell} \setminus B_R(0)$. Then for $x \in M_j$

$$\int_{\eta_{x,\lambda}(M)} f\, d\mathcal{H}^n = \int_{\eta_{x,\lambda}(N_j)} f\, d\mathcal{H}^n - \int_{\eta_{x,\lambda}(N_j \setminus M_j)} f\, d\mathcal{H}^n + \int_{\eta_{x,\lambda}(M \setminus M_j)} f\, d\mathcal{H}^n$$

and, since $x \in N_j$ and N_j is a C^1 submanifold,

$$\lim_{\lambda \downarrow 0} \int_{\eta_{x,\lambda}(N_j)} f\, d\mathcal{H}^n = \int_{T_x N_j} f\, d\mathcal{H}^n,$$

and, by the Upper density Theorem 3.6 of Ch. 1,

$$\begin{aligned}\left|\int_{\eta_{x,\lambda}(M \setminus N_j)} f\, d\mathcal{H}^n\right| &\leq \sup|f| \mathcal{H}^n(B_R(0) \cap \eta_{x,\lambda}(M \setminus N_j)) \\ &= \sup|f| \lambda^{-n} \mathcal{H}^n(B_{\lambda R}(x) \cap M \setminus N_j) \to 0 \text{ for } \mathcal{H}^n\text{-a.e. } x \in M_j\end{aligned}$$

Similarly, again by the Upper density Theorem,

$$\Big|\int_{\eta_{x,\lambda}(N_j\setminus M_j)} f\,d\mathcal{H}^n\Big| \to 0 \text{ for } \mathcal{H}^n\text{-a.e. } x \in M_j.$$

Thus we have shown that T_xM exists and $= T_xN_j$ for $\mathcal{H}^n$a.e. $x \in M_j$. In particular Remark 1.5(2) is checked. □

Proof of 1.6 "⇐": Define $\mu = \mathcal{H}^n \llcorner (M \cap B_R(0))$, we have that μ is Borel regular with $\mu(\mathbb{R}^{n+\ell}) < \infty$.

Given any ℓ-dimensional subspace $\pi \subset \mathbb{R}^{n+\ell}$ and any $\alpha \in (0,1)$ we let $X_\alpha(\pi, x)$ denote the cone

$$X_\alpha(\pi,x) = \{y \in \mathbb{R}^{n+\ell} : \operatorname{dist}(y-x,\pi) \le \alpha|y-x|\}, \tag{1}$$

which can alternatively be written

$$X_\alpha(\pi,x) = \{y \in \mathbb{R}^{n+\ell} : |q_\pi(y-x)| \le \alpha|y-x|\}, \tag{2}$$

where q_π denotes orthogonal projection of $\mathbb{R}^{n+\ell}$ onto $\pi^\perp$, with

$$\pi^\perp = \{z \in \mathbb{R}^{n+\ell} : z\cdot w = 0\,\forall w \in \pi\}.$$

For ℓ-dimensional subspaces π, π' we define the distance between π, π', denoted $d(\pi,\pi')$, by

$$d(\pi,\pi') = \sup_{|x|=1} |q_\pi(x) - q_{\pi'}(x)|, \tag{3}$$

so that in fact $d(\pi,\pi')$ is just the norm $\|q_\pi - q_{\pi'}\|$ of the linear map $q_\pi - q_{\pi'}$. Since T_xM exists μ-a.e. and $\mu(\mathbb{R}^{n+\ell}) < \infty$, by 1.15 of Ch. 1 we can choose a closed subset $F \subset M$ such that

$$\mu(\mathbb{R}^{n+\ell} \setminus F) \le \frac{1}{4}\mu(\mathbb{R}^{n+\ell}) \tag{4}$$

and such that for each $x \in F$, M has an approximate tangent space P_x at x. Thus in particular by Remark 1.5(3) we have

$$\lim_{\rho\downarrow 0} \frac{\mu(B_\rho(x))}{\omega_n\rho^n} = 1 \tag{5}$$

and

$$\lim_{\rho\downarrow 0} \frac{\mu(X_{\frac12}(\pi_x,x) \cap B_\rho(x))}{\omega_n\rho^n} = 0, \tag{6}$$

for all $x \in F$, where $\pi_x = (P_x)^\perp$.

For $k = 1, 2, \ldots$ and $x \in F$, define

$$f_k(x) = \inf_{0<\rho<\frac1k} \frac{\mu(B_\rho(x))}{\omega_n\rho^n}$$

and

$$q_k(x) = \sup_{0<\rho<\frac1k} \frac{\mu(X_{\frac12}(\pi_x,x) \cap B_\rho(x))}{\omega_n\rho^n}.$$

Then

$$\lim f_k(x) = 1 \text{ and } \lim q_k(x) = 0 \ \forall x \in F, \tag{7}$$

and hence by Egoroff's Theorem we can choose a closed set $E \subset F$ with

$$\mu(F \setminus E) \le \frac14 \mu(\mathbb{R}^{n+\ell}) \tag{8}$$

and with (7) holding *uniformly* for $x \in E$. Thus for each $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\frac{\mu(B_\rho(x))}{\omega_n\rho^n} \ge 1-\varepsilon, \quad \frac{\mu(X_{\frac12}(\pi_x,x) \cap B_\rho(x))}{\omega_n\rho^n} \le \varepsilon \tag{9}$$

$x \in E$, $0 < \rho < \delta$.

Now choose ℓ-dimensional subspaces $\pi_1, \ldots, \pi_N$ of $\mathbb{R}^{n+\ell}$ ($N = N(n,\ell)$) such that for *each* ℓ-dimensional subspace π of $\mathbb{R}^{n+\ell}$, there is a $j \in \{1,\ldots,N\}$ such that $d(\pi,\pi_j) < \frac{1}{16}$, and let $E_1, \ldots, E_N$ be the subsets of E defined by

$$E_j = \{x \in E : d(\pi_j, \pi_x) < \tfrac{1}{16}\}.$$

Then $E = \cup_{j=1}^N E_j$ and we claim that if we take $\varepsilon = 1/16^n$ and let $\delta > 0$ be such that (9) holds, then

$$X_{\frac14}(\pi_j,x) \cap E_j \cap B_{\delta/2}(x) = \{x\}, \ \forall x \in E_j, \ j = 1,\ldots,N. \tag{10}$$

Indeed otherwise we could find a point $x \in E_j$ and a $y \in X_{\frac14}(\pi_j,x) \cap E_j \cap \partial B_\rho(x)$ for some $0 < \rho \le \delta/2$. But since $x \in E$ and $2\rho \le \delta$, we have (by (9))

$$\mu(X_{\frac12}(\pi_x,x) \cap B_{2\rho}(x)) < \varepsilon\omega_n(2\rho)^n$$

and (since $B_{\rho/8}(y) \subset X_{\frac12}(\pi_j,x) \cap B_{2\rho}(x)$) we have also (again by (9))

$$\begin{aligned}\mu(X_{\frac12}(\pi_x,x) \cap B_{2\rho}(x)) &\ge \mu(B_{\rho/8}(y))\\ &\ge \omega_n(\rho/8)^n,\end{aligned}$$

which contradicts (9), since $\varepsilon = 1/16^n$. We have therefore proved (10). Now for any fixed $x_0 \in E_j$ it is easy to check that (10), taken together with the extension theorem 1.2 of Ch. 2, implies

$$E_j \cap B_{\delta/4}(x_0) \subset Q^T(\text{graph } f)$$

where Q is an orthogonal transformation of $\mathbb{R}^{n+\ell}$ with $Q(\pi_j) = \{0\} \times \mathbb{R}^\ell$, and where $f = (f^1, \ldots, f^\ell) : \mathbb{R}^n \to \mathbb{R}^\ell$ is Lipschitz.

Since $j \in \{1, \ldots, N\}$ and $x_0 \in E_j$ are arbitrary, we can then evidently select Lipschitz functions $f_1, \ldots, f_J : \mathbb{R}^n \to \mathbb{R}^\ell$ and orthogonal transformations $Q_1, \ldots, Q_J$ of $\mathbb{R}^{n+\ell}$ such that

$$E \subset \cup_{j=1}^J Q_j(\text{graph } f_j).$$

Thus by (4), (8) we have

$$\mu(\mathbb{R}^{n+\ell} \setminus \cup_{j=1}^J Q_j(\text{graph } f_j)) \le \frac{1}{2}\mu(\mathbb{R}^{n+\ell}).$$

Since we can now repeat the argument, starting with $\mu \llcorner (\mathbb{R}^{n+\ell} \setminus \cup_{j=1}^J Q_j(\text{graph } f_j))$ in place of μ, we thus deduce that there are countably many Lipschitz graphs graph f_j, $j = 1, 2, \ldots$, $f_j : \mathbb{R}^n \to \mathbb{R}^\ell$, and corresponding orthogonal transformations $Q_1, Q_2, \ldots$ with that $\mu(\mathbb{R}^{n+\ell} \setminus \cup_{j=1}^\infty Q_m(\text{graph } f_j)) = 0$. Taking G_j to be the graph map $x \mapsto (x, f_j(x))$ and $F_j = Q_j \circ G_j$ we then have that $F_j : \mathbb{R}^n \to \mathbb{R}^{n+\ell}$ are Lipschitz and $\mathcal{H}^n(M \setminus (\cup_{j=1}^\infty F_j(\mathbb{R}^n))) = 0$, so M is countably n-rectifiable. □

It is often convenient to be able to relax the condition $\mathcal{H}^n(M \cap K) < \infty\ \forall$ compact K in 1.4 and 1.6 and consider instead sets M which can be written as the countable union $\cup_{j=1}^\infty M_j$ of $\mathcal{H}^n$-measurable sets M_j with $\mathcal{H}^n(M_j \cap K) < \infty$ for each j and each compact $K \subset \mathbb{R}^{n+\ell}$. This is evidently equivalent to the requirement that M is $\mathcal{H}^n$-measurable and there exists a positive $\mathcal{H}^n$-measurable function θ on M such that $\int_{M \cap K} \theta \, d\mathcal{H}^n < \infty$ for each compact K, so we proceed to discuss this situation, starting with the definition of approximate tangent space in such a setting:

1.7 Definition: Let M be an $\mathcal{H}^n$-measurable $\mathbb{R}^{n+\ell}$ and let θ be a positive $\mathcal{H}^n$-measurable function on M with $\int_{M \cap K} \theta \, d\mathcal{H}^n < \infty$ for each compact $K \subset \mathbb{R}^{n+\ell}$. For each $x \in \mathbb{R}^{n+\ell}$, we say an n-dimensional subpace P_x is an approximate tangent space of M with respect to θ if

$$(\ddagger) \qquad \lim_{\lambda \downarrow 0} \int_{\eta_{x,\lambda}(M)} f(y)\theta(x + \lambda y) \, d\mathcal{H}^n(y) = \theta(x) \int_{P_x} f(y) \, d\mathcal{H}^n(y)$$

for each $f \in C_c^0(\mathbb{R}^{n+\ell})$. Evidently P_x is unique if it exists at all so we denote it $T_x M$, and also $T_x M$ agrees with our previous notion of approximate tangent space in case $\mathcal{H}^n(M \cap K) < \infty$ for all compact $K \subset \mathbb{R}^{n+\ell}$ and $\theta \equiv 1$.

1.8 Remarks: (1) Notice that if M, θ are as in the above definition, then, by Lusin's Theorem 1.17 of Ch. 1, there is an increasing sequence $\{M_j\}$ of compact subsets of M with $\mathcal{H}^n(M \setminus (\cup_{j=1}^\infty M_j)) = 0$ and $\theta | M_j$ continuous (hence has a positive lower bound, hence $\mathcal{H}^n(M_j) < \infty$) for each j. Also, since $\mu = \mathcal{H}^n \llcorner \theta$ is locally finite in $\mathbb{R}^{n+\ell}$ we can apply Theorem 3.6 of Ch. 1 to give a set $E_j \subset \mathbb{R}^{n+\ell}$ of $\mathcal{H}^n$-measure zero such that $\Theta^{*n}(\mu \llcorner (M \setminus M_j), x) = 0$ for each j and each $x \in M_j \setminus E_j$. Then it is straightforward to check that, with $E = \cup_{j=1}^\infty E_j$, for every j and every $x \in M_j \setminus E$, $T_x M_j$ exists as in Definition 1.4 $\iff$ $T_x M$ exists as in Definition 1.7, and then $T_x M = T_x M_j$.

(2) By taking a C^0 function $f : \mathbb{R}^{n+\ell} \to [0,1]$ with $f \equiv 1$ in $B_1(0)$ and $f \equiv 0$ in $\mathbb{R}^{n+\ell} \setminus B_{1+\varepsilon}(0)$, we see that the definition ($\ddagger$) implies in particular that

$$\lim_{\rho \downarrow 0} (\omega_n \rho^n)^{-1} \int_{M \cap B_\rho(y)} \theta \, d\mathcal{H}^n = \theta(y)$$

whenever M has an approximate tangent space with respect to θ.

In view of the Remark 1.8 (1) above, we can apply Theorem 1.6 to the sets M_j to deduce the following generalization of Theorem 1.6:

1.9 Theorem. *Suppose $M \subset \mathbb{R}^{n+\ell}$ is $\mathcal{H}^n$-measurable and θ is a positive $\mathcal{H}^n$-measurable function M with $\int_{M \cap K} \theta \, d\mathcal{H}^n < \infty$ for each compact $K \subset \mathbb{R}^{n+\ell}$. Then M is countably n-rectifiable $\iff$ M has an approximate tangent space $T_x M$ with respect to θ for $\mathcal{H}^n$-a.e. $x \in M$.*

2 Gradients, Jacobians, Area, Co-Area

Throughout this section M is supposed to be $\mathcal{H}^n$-measurable and countably n-rectifiable, so that we can express M as the disjoint union $\cup_{j=0}^\infty M_j$ (as in Remark 1.3 of the present chapter), where $\mathcal{H}^n(M_0) = 0$, M_j is $\mathcal{H}^n$-measurable of finite $\mathcal{H}^n$-measure, and $M_j \subset N_j$, $j \ge 1$, where N_j are embedded n-dimensional C^1 submanifolds of $\mathbb{R}^{n+\ell}$.

Let f be a locally Lipschitz function on U, where U is an open set in $\mathbb{R}^{n+\ell}$ containing M. Then according to the discussion in §4 of Ch. 2 we can define the gradient of f, $\nabla^M f$, $\mathcal{H}^n$-a.e. $y \in M$ by

2.1 Definition:

$$\nabla^M f(y) = \nabla^{N_j} f(y), \quad y \in M_j,$$

where the notation is as above.

Notice that, up to change on sets of $\mathcal{H}^n$-measure zero, this is independent of the decomposition $M = \cup M_j$ (and independent of the choice of the C^1 submanifolds N_j). Because for $\mathcal{H}^n$-a.e. $y \in M$ we have $D_\tau f(y) = \frac{d}{dt} f(y+t\tau)|_{\tau=0}$ for all $\tau \in T_x M$ by 4.10 of Ch. 2 and the fact that $T_y M$ agrees with $T_y N_j$ for $\mathcal{H}^n$-a.e. $y \in M_j (= M \cap N_j)$. So $\nabla^M f$ is well defined as an L^1 function with respect to Hausdorff measure $\mathcal{H}^n$ on M.

Having defined $\nabla^M f$, we can now define the linear $d^M f_x : T_x M \to \mathbb{R}$ induced by f by setting

$$d^M f_x(\tau) = D_\tau f(y) \left(= \langle \tau, \nabla^M f(x)\rangle\right), \ \tau \in T_x M$$

at all points where $T_x M$ and $\nabla^M f(x)$ exist. More generally, if $f = (f^1, \ldots, f^Q)$ takes values in $\mathbb{R}^Q$ (f^j still locally Lipschitz on U, $j = 1, \ldots, \ell$), we define $d^M f_x : T_x M \to \mathbb{R}^Q$ by

2.2 $$d^M f_x(\tau) = D_\tau f(x).$$

With such an f, in case $Q = n + \ell_1$ ($\ell_1 \geq 0$), we define the Jacobian $J_M f(x)$ for $\mathcal{H}^n$-a.e. $x \in M$ as in 4.12 of Ch. 2; thus

2.3 $$J_f^M(x) = \sqrt{\det \mathcal{J}(x)} = \sqrt{\det\left(d^M f_x\right)^* \circ \left(d^M f_x\right)}$$

where $\mathcal{J}(x)$ is the matrix with $D_{\tau_p} f(x) \cdot D_{\tau_q} f(x)$ in the p-th row and q-th column ($\tau_1, \ldots, \tau_n$ any orthonormal basis for $T_x M$) and $\left(d^M f_x\right)^* : \mathbb{R}^{n+\ell_1} \to T_x M$ denotes the adjoint of $d^M f_x$. In view of 4.11 and the area formula for Lipschitz maps from domains in $\mathbb{R}^n$, as discussed in §3 we have the general area formula

2.4 $$\int_A J_f^M \, d\mathcal{H}^n = \int_{\mathbb{R}^{n+\ell}} \mathcal{H}^0\left(A \cap f^{-1}(y)\right) d\mathcal{H}^n(y)$$

for any $\mathcal{H}^n$-measurable set $A \subset M$. The proof of this is as follows:

We may suppose (decomposing $\mathcal{H}^n$-almost all M_j as a countable union if necessary and using the C^1 Approximation Theorem 1.5 of Ch. 2) that $f|M_j = g_j|M_j$, where g_j is C^1 on $\mathbb{R}^{n+\ell}$, $j \geq 1$.

By virtue of the 2.1, 2.2, we then have

$$J_f^M(x) = J^{N_j} g_j(x), \ \mathcal{H}^n\text{-a.e. } x \in M_j.$$

Thus J_f^M is $\mathcal{H}^n$-measurable, and by the smooth case 3.4 of Ch. 2 of the area formula (with N_j in place of M, $A \cap M_j$ in place A and g_j in place of f), we have

$$\int_{A \cap M_j} J_f^M \, d\mathcal{H}^n = \int_{\mathbb{R}^n} \mathcal{H}^0\left(A \cap M_j \cap f^{-1}(y)\right) d\mathcal{H}^n.$$

We now conclude 2.4 by summing over $j \geq 1$ and using the (easily checked) fact that if $\psi : U \to \mathbb{R}^m$ is locally Lipschitz and B has $\mathcal{H}^n$-measure zero, then $\mathcal{H}^n(\psi(B)) = 0$.

We note also that if h is any non-negative $\mathcal{H}^n$-measurable function on M, then, by approximation of h by simple functions, 2.4 implies the more general formula

2.5 $$\int_M h \, J_f^M \, d\mathcal{H}^n = \int_{\mathbb{R}^{n+\ell}} \Big(\int_{f^{-1}(y)} h \, d\mathcal{H}^0\Big) \, d\mathcal{H}^n(y).$$

In case $f|M$ is 1:1 and this becomes

2.6 $$\int_M h \, J_f^M \, d\mathcal{H}^n = \int_{f(M)} h \circ f^{-1} \, d\mathcal{H}^n.$$

There is also a version of the co-area formula in case M is merely $\mathcal{H}^n$-measurable, countably n-rectifiable and $f : U \to \mathbb{R}^m$ is locally Lipschitz with $m < n$; we write $n = m + k$ and here U open with $M \subset U$ as before.

In fact we can define (Cf. the smooth case described in §6 of Ch. 2)

2.7 $$J_f^{M*}(x) = \sqrt{\det\left(d^M f_x\right) \circ \left(d^M f_x\right)^*}$$

with $d^M f_x$ as in 2.2 and $\left(d^M f_x\right)^*$ = adjoint of $d^M f_x$. Then, for any $\mathcal{H}^n$-measurable set $A \subset M$,

2.8 $$\int_A J_f^{M*} \, d\mathcal{H}^n = \int_{\mathbb{R}^m} \mathcal{H}^k\left(A \cap f^{-1}(y)\right) d\mathcal{L}^m(y).$$

This follows from the C^1 case (see §6 of Ch. 2) by using the decomposition $M = \cup_{j=0}^\infty M_j$ of Remark 1.3 and the C^1 Approximation Theorem 1.5 of Ch. 2 in a similar manner to the procedure used for the discussion of the area formula above.

As for the smooth case, approximating a given non-negative $\mathcal{H}^n$-measurable function g by simple functions, we deduce directly from 2.8 the more general formula

2.9 $$\int_A g \, J_f^{M*} \, d\mathcal{H}^n = \int_{\mathbb{R}^m} \Big(\int_{f^{-1}(y) \cap M} g \, d\mathcal{H}^k\Big) d\mathcal{L}^m(y).$$

2.10 Remarks: (1) Note that Remark 6.7 of Ch. 2 carries over without change to this setting.

(2) The "slices" $M \cap f^{-1}(y)$ are countably k-rectifiable subsets of $\mathbb{R}^{n+\ell}$ for $\mathcal{L}^m$-a.e. $y \in \mathbb{R}^m$. This follows directly from the decomposition $M = \cup_{j=0}^\infty M_j$ of Remark 1.3 together with the C^1 Sard-type Theorem 6.4 of Ch. 2 and the C^1 Approximation Theorem 1.5 of Ch. 2.

3 Purely Unrectifiable Sets, Structure Theorem

3.1 Definition: A subset $S \subset \mathbb{R}^{n+\ell}$ is said to be purely n-unrectifiable if P contains no countably n-rectifiable subsets of positive $\mathcal{H}^n$-measure.

3.2 Lemma. *If A is an arbitrary "$\mathcal{H}^n$ σ-finite" subset of $\mathbb{R}^{n+\ell}$ (i.e. $A = \cup_{j=1}^{\infty} A_j$ with $\mathcal{H}^n(A_j) < \infty$ for each j), it is always possible to decompose A into a disjoint union*

$$A = R \cup P,$$

where R is countably n-rectifiable and P is purely n-unrectifiable. Also R can be chosen to be a Borel set if A is $\mathcal{H}^n$ measurable.

Proof: First observe that in case in case A is $\mathcal{H}^n$-measurable we can also take each A_j to be $\mathcal{H}^n$-measurable (e.g., first take a Borel set $B_j \supset A_j$ with $\mathcal{H}^n(B_j) = \mathcal{H}^n(A_j)$ and then replace A_j by $A \cap B_j$). In this case we let

$$\alpha_j = \sup\{\mathcal{H}^n(S) : S \subset A_j,\ S \text{ countably } n\text{-rectifiable, } \mathcal{H}^n\text{-measurable}\}.$$

By 1.15(2) of Ch. 1 and the definition of α_j we can select closed countably n-rectifiable sets $R_{ij} \subset A_j$ with $\mathcal{H}^n(R_{ij}) > \alpha_j - \frac{1}{i}$ and let $R = \cup_{i,j} R_{ij}$. Evidently R is a countably n-rectifiable Borel set, $A \setminus R$ is purely unrectifiable, and 3.2 is proved with R Borel and P $\mathcal{H}^n$-measurable.

To handle the case when A is not necessarily $\mathcal{H}^n$-measurable, we first pick a Borel set $B = \cup_j B_j$, where each B_j is a Borel set containing A_j with the same $\mathcal{H}^n$-measure as A_j. Then by the measurable case of 3.2 we have $B = R \cup P$ with R countably n-rectifiable Borel and P purely unrectifiable, and then $A = (A \cap R) \cup (A \cap P)$ is a suitable decomposition of A. $\square$

The following lemma gives a simple and convenient sufficient condition for checking if a set is purely n-unrectifiable. In this lemma we adopt the notation that p_L denotes the orthogonal projection of $\mathbb{R}^{n+\ell}$ onto L for any n-dimensional subspace $L \subset \mathbb{R}^{n+\ell}$.

3.3 Lemma. *For $1 \le j_1 < j_2 < \cdots < j_n \le n+\ell$ let $p_{j_1,\ldots,j_n}$ denote the orthogonal projection of $\mathbb{R}^{n+\ell}$ onto $\operatorname{span}\{e_{j_1},\ldots,e_{j_n}\}$, and suppose $S \subset \mathbb{R}^{n+\ell}$ has the property that $\mathcal{H}^n(p_{j_1,\ldots,j_n}(S)) = 0$ for each $1 \le j_1 < \cdots < j_n \le n+\ell$. Then S is purely n-unrectifiable.*

Proof of 3.3: We first claim that if L is any n-dimensional subspace of $\mathbb{R}^{n+\ell}$ then there is some $1 \le j_1 < j_2 < \cdots < j_n \le n+\ell$ such that $p_{j_1,\ldots,j_n}|L$ is an isomorphism onto $\operatorname{span}\{e_{j_1},\ldots,e_{j_n}\}$. Indeed if $v_1,\ldots,v_n$ is a basis for L and V is the $n \times (n+\ell)$ matrix with rows $v_1,\ldots,v_n$, then there is $1 \le j_1 < \cdots < j_n \le n+\ell$ such that the column numbers $j_1,\ldots,j_n$ of V give an $n \times n$ non-singular matrix, and hence $p_{j_1,\ldots,j_n}(v_1),\ldots,p_{j_1,\ldots,j_n}(v_n)$ are linearly independent.

Now suppose on the contrary that S is not purely n-unrectifiable. Then Lemma 1.2 implies there is an n-dimensional C^1 submanifold N with $\mathcal{H}^n(S \cap N) > 0$, so there must be some $x \in S \cap N$ with $\mathcal{H}^n(S \cap N \cap B_\rho(x)) > 0$ for all $\rho > 0$. With such an x we see that, by the above discussion and by Remark 4.4 of Ch. 2 (with $M = N$) that there is $1 \le j_1 < j_2 < \cdots < j_n \le n+\ell$ and $\rho > 0$ such that $p_{j_1,\ldots,j_n}|N \cap \breve{B}_\rho(x)$ is a C^1 diffeomorphism onto an open $W \subset \operatorname{span}\{e_{j_1},\ldots,e_{j_n}\}$ and so $\mathcal{H}^n(p_{j_1,\ldots,j_n}(S \cap N \cap B_\rho(x))) > 0$. $\square$

3.4 Remark. The above proof can be modified to give a more general conclusion: If $v_1,\ldots,v_{n+\ell}$ is any basis for $\mathbb{R}^{n+\ell}$, if $p_{v_{j_1},\ldots,v_{j_n}}$ denotes orthogonal projection of $\mathbb{R}^{n+\ell}$ onto $\operatorname{span}\{v_{j_1},\ldots,v_{j_n}\}$, and if $S \subset \mathbb{R}^{n+\ell}$ is such that $\mathcal{H}^n(p_{v_{j_1},\ldots,v_{j_n}}(S)) = 0$ for each $1 \le j_1 < j_2 < \cdots < j_n \le n+\ell$, then S is purely n-unrectifiable.

3.5 Example. A simple example (in the case $n = \ell = 1$) of the use of Lemma 3.3 is the following: Let $C_0 = [0,1] \times [0,1]$, $C_1 =$ the union of the 4 sub-squares of C_0 with edge length $\frac{1}{4}$ each sharing one corner with C_0. Observe that the orthogonal projection p onto the line $y = \frac{1}{2}x$ projects C_1 onto a full line segment σ of length $\frac{3}{\sqrt{5}}$. Thus if we inductively define a sequence C_n of sets, each of which is the union of 4^n squares with edge length 4^{-n} and if we stipulate that C_{n+1} is obtained from C_n by replacing each square s of C_n with 4 squares of edge-length 4^{-n-1}, each sharing a corner with s, then $C_{n+1} \subset C_n$ and each C_n projects via the orthogonal projection p onto the full line segment σ, and hence so does the compact set $C = \cap_{n=0}^{\infty} C_n$. Furthermore $\mathcal{H}^1(C) \ge \mathcal{H}^1(p(C)) = \frac{3}{\sqrt{5}} > 0$, and also $\mathcal{H}^1(C) \le \sqrt{2} < \infty$ because each of the 4^n squares comprising C_n has diameter $4^{-n}\sqrt{2}$. Finally, each C_n projects via orthogonal projection p_x of $\mathbb{R}^2$ onto the x-axis to a union of 2^n closed intervals each of length 4^{-n}, and hence $\mathcal{L}^1(p_x(C)) = \lim \mathcal{L}^1(p_x(C_n)) = 0$. Similarly $\mathcal{L}^1(p_y(C)) = 0$, where p_y denotes orthogonal projection onto the y-axis. Evidently then Lemma 3.3 is applicable with $n = \ell = 1$, so C is purely 1-unrectifiable.

A very non-trivial theorem (the Structure Theorem) due to Besicovitch [Bes28, Bes38, Bes39] in case $n = \ell = 1$ and Federer [Fed69] in general, says that the purely unrectifiable sets Q of $\mathbb{R}^{n+\ell}$ which (like the subset P in 3.2) can be written as the countable union of sets of finite $\mathcal{H}^n$-measure, are characterized by the fact that they have $\mathcal{H}^n$-null projection via almost all orthogonal projections onto n-dimensional subspaces of $\mathbb{R}^{n+\ell}$. More precisely:

3.6 Theorem. *Suppose Q is a purely n-unrectifiable subset of $\mathbb{R}^{n+\ell}$ with $Q = \cup_{j=1}^{\infty} Q_j$, $\mathcal{H}^n(Q_j) < \infty\ \forall j$. Then $\mathcal{H}^n(p(Q)) = 0$ for σ-almost all $p \in O(n+\ell, n)$. Here σ is Haar measure for $O(n+\ell, n)$, the orthogonal projections of $\mathbb{R}^{n+\ell}$ onto n-dimensional subspaces of $\mathbb{R}^{n+\ell}$.*

For the proof of this theorem see [Fed69] or [Ros84].

3.7 Remark: Of course only the purely n-unrectifiable subsets could possibly have the null projection property described in 3.6, by virtue of Lemma 3.3 above.

Notice that, by combining 3.2 and 3.7, we get the following *Rectifiability Theorem*, which is of fundamental importance in understanding the structure of subsets of $\mathbb{R}^{n+\ell}$:

3.8 Theorem (Rectifiability Theorem for sets.) *If A is an arbitrary subset of $\mathbb{R}^{n+\ell}$ which can be written as a countable union $\cup_{j=1}^{\infty} A_j$ with $\mathcal{H}^n(A_j) < \infty\ \forall j$, and if every subset $B \subset A$ with positive $\mathcal{H}^n$-measure has the property that $\mathcal{H}^n(p(B)) > 0$ for a set of $p \in O(n+\ell, n)$ with σ-measure> 0, then A is countably n-rectifiable.*

4 Sets of Locally Finite Perimeter

An important class of countably n-rectifiable sets in $\mathbb{R}^{n+\ell}$ comes from the sets of locally finite perimeter. (Or Cacciopoli sets—see De Giorgi [DG61], Giusti [Giu84].) First we need some definitions.

If $U \subset \mathbb{R}^{n+1}$ is open and if E is an $\mathcal{L}^{n+1}$-measurable subset of $\mathbb{R}^{n+1}$, we say that E has locally finite perimeter in U if the characteristic function χ_E of E is in $BV_{\mathrm{loc}}(U)$. (See §2 of Ch. 2.)

Thus E has locally finite perimeter in U if there is a Radon measure μ_E $(= |D\chi_E|$ in the notation of §2 of Ch. 2) on U and a μ_E-measurable function $\nu = (\nu^1, \ldots, \nu^{n+1})$ with $|\nu| = 1$ μ_E-a.e. in U, such that

$$\textbf{4.1} \qquad \int_{E \cap U} \operatorname{div} g \, d\mathcal{L}^{n+1} = -\int_U g \cdot \nu \, d\mu_E$$

for each $g = (g^1, \ldots, g^{n+1})$ with $g^j \in C_c^1(U)$, $j = 1, \ldots, n+1$. Notice that if E is open and $\partial E \cap U$ is an n-dimensional embedded C^1 submanifold of $\mathbb{R}^{n+1}$, then the divergence theorem tells us that 4.1 holds with $\mu_E = \mathcal{H}^n \llcorner (\partial E \cap U)$ and with $\nu =$ the inward pointing unit normal to ∂E. Thus in general we interpret μ_E as a "generalized boundary measure" and ν as a "generalized inward unit normal". It turns out (see 4.3 below) that in fact this interpretation is quite generally correct in a rather precise (and concrete) sense.

We now want to define the *reduced boundary* $\partial^* E$ of a set E of finite perimeter by

$$\textbf{4.2} \qquad \partial^* E = \Big\{x \in U : \lim_{\rho \downarrow 0} \frac{\int_{B_\rho(x)} \nu \, d\mu_E}{\mu_E(B_\rho(x))} \text{ exists and has length } 1\Big\}.$$

Since $|\nu| = 1$ μ_E-a.e. in U, by virtue of XXX Theorem 3.23 of Ch. 1 we have $\mu_E(U \setminus \partial^* E) = 0$, so that $\mu_E = \mu_E \llcorner \partial^* E$. We in fact claim much more:

4.3 Theorem (De Giorgi.) *Suppose E has locally finite perimeter in U. Then $\partial^* E$ is countably n-rectifiable and $\mu_E = \mathcal{H}^n \llcorner \partial^* E$. In fact at each point $x \in \partial^* E$ the approximate tangent space T_x of μ_E exists, has multiplicity 1, and is given by*

$$(\ddagger) \qquad T_x = \{y \in \mathbb{R}^{n+1} : y \cdot \nu_E(x) = 0\},$$

where $\nu_E(x) = \lim_{\rho \downarrow 0} \frac{\int_{B_\rho(x)} \nu \, d\mu_E}{\mu_E(B_\rho(x))}$ (so that $|\nu_E(x)| = 1$ by 4.2). Furthermore at any such point $x \in \partial^* E$ we have that $\nu_E(x)$ is the "inward pointing unit normal for E" in the sense that

$$(\ddagger\ddagger) \qquad E_{x,\lambda} \equiv \{\lambda^{-1}(y - x) : y \in E\} \to \{y \in \mathbb{R}^{n+1} : y \cdot \nu_E(x) > 0\}$$

in the $L^1_{\mathrm{loc}}(\mathbb{R}^{n+1})$ sense.

Proof: By 1.6 and 3.6 of Ch. 1, the first part of the theorem follows from 4.3 (‡), which we now establish. 4.3 (‡‡) will also appear as a "by product" of the proof of 4.3 (‡). Assume without loss of generality $\nu \equiv \nu_E$ on $\partial^* E$.

Take any $y \in \partial^* E$. For convenience of notation we suppose that $y = 0$ and $\nu(0) = (0, \ldots, 0, 1)$. Then we have

$$(1) \qquad \lim_{\rho \downarrow 0} \frac{\int_{B_\rho(0)} \nu_{n+1} \, d\mu_E}{\mu_E(B_\rho(0))} = 1$$

and hence (since $|\nu| = 1$ μ_E-a.e.)

$$(2) \qquad \lim_{\rho \downarrow 0} \frac{\int_{B_\rho(0)} |\nu_i| \, d\mu_E}{\mu_E(B_\rho(0))} = 0, \ i = 1, \ldots, n.$$

further if $\zeta \in C_0^1(U)$ has support in $B_\rho(0) \subset U$, then by 4.1

$$(3) \qquad \begin{aligned} \int_U \nu_{n+1} \zeta \, d\mu_E &= -\int_U \chi_E D_{n+1} \zeta \, d\mathcal{L}^{n+1} \\ &\le \int_E |D\zeta| \, d\mathcal{L}^{n+1}. \end{aligned}$$

Now (taking $B_\rho(0)$ to be the closed ball) replace ζ by a decreasing sequence $\{\zeta_k\}$ converging pointwise to the characteristic function of $B_\rho(0)$ and satisfying

$$\lim_{k\to\infty}\int_E |D\zeta_k| = \frac{d}{d\rho}\mathcal{L}^{n+1}(E\cap B_\rho(0)). \tag{4}$$

(Notice that this can be done whenever the right side exists, which is $\mathcal{L}^1$-a.e. ρ.) Then (3) gives

$$\int_{B_\rho(0)} \nu_{n+1}\,d\mu_E \le \frac{d}{d\rho}\mathcal{L}^{n+1}(E\cap B_\rho(0)) \tag{5}$$

for $\mathcal{L}^1$-a.e. $\rho\in(0,\rho_0)$, $\rho_0 = \operatorname{dist} 0,\partial U$. Then by (1) we have, for suitable $\rho_1\in(0,\rho_0)$,

$$\begin{aligned}\mu_E(B_\rho(0)) &\le 2\frac{d}{d\rho}\mathcal{L}^{n+1}(E\cap B_\rho(0)) \equiv 2\mathcal{H}^n(E\cap\partial B_\rho(0))\\ &\le 2(n+1)\omega_{n+1}\rho^n\end{aligned} \tag{6}$$

for $\mathcal{L}^1$-a.e. $\rho\in(0,\rho_1)$.

Then by the Compactness Theorem 2.6 of Ch. 2, it follows that we can select a sequence $\rho_k\downarrow 0$ so that $\chi_{\rho_k^{-1}E}\to\chi_R$ in $L^1_{\text{loc}}(\mathbb{R}^{n+1})$, where F is a set of locally finite perimeter in $\mathbb{R}^{n+1}$. Hence in particular for any non-negative $\zeta\in C^1_0(\mathbb{R}^{n+1})$

$$\lim_{k\to\infty}\int_{\rho_k^{-1}E} D_i\zeta\,d\mathcal{L}^{n+1} = \int_F D_i\zeta\,d\mathcal{L}^{n+1}. \tag{7}$$

Now write $\zeta_k(x)=\zeta(\rho_k^{-1}x)$ and change variable $x\to\rho_k x$; then

$$\int_{\rho_k^{-1}E} D_i\zeta\,d\mathcal{L}^{n+1} = \rho_k^{-n}\int_E D_i\zeta_k\,d\mathcal{L}^{n+1} \equiv -\rho_k^{-n}\int_U \zeta_k\nu_i\,d\mu_E \tag{8}$$

(by 4.1), so that $\int_{\rho_k^{-1}E} D_i\zeta\,d\mathcal{L}^{n+1}\to 0$ by (2) for $i=1,\ldots,n$. Thus (7) gives

$$\int_F D_i\zeta\,d\mathcal{L}^{n+1} = 0\ \forall\zeta\in C^1_0(\mathbb{R}^{n+1}),\ i=1,\ldots,n,$$

and it follows that $F=\mathbb{R}^n\times H$ for some $\mathcal{L}^1$-measurable subset H of $\mathbb{R}$.

On the other hand by 4.1 with $g=\zeta_k e_{n+1}$ and by (1) we have, for k sufficiently large and $\zeta\ge 0$,

$$\begin{aligned}0 \le \rho_k^{-n}\int_U \zeta_k\nu_{n+1}\,d\mu_E &= \int_{\rho_k^{-1}E} D_{n+1}\zeta\\ \to \int_F D_{n+1}\zeta &\equiv \int_{\mathbb{R}^n}\Big(\int_H \frac{\partial\zeta}{\partial x^{n+1}}(x',x^{n+1})\,dx^{n+1}\Big)\,dx'\end{aligned}$$

as $k\to\infty$, so that χ_H is *non-decreasing* on $\mathbb{R}$, hence

$$F = \{x\in\mathbb{R}^{n+1} : x^{n+1} < \lambda\} \tag{9}$$

for some λ. We have next to show that $\lambda = 0$. To check this we use the Sobolev inequality (see e.g. [GT01]) to deduce that, if $\zeta\ge 0$, $\operatorname{spt}\zeta\subset U$ and $\sigma<\operatorname{dist}(\operatorname{spt}\zeta,\partial U)$, then

$$\begin{aligned}&\Big(\int_U (\zeta\varphi_\sigma * \chi_E)^{\frac{n+1}{n}}\,d\mathcal{L}^{n+1}\Big)^{\frac{n}{n+1}}\\ &\qquad\le c\int_U |D(\zeta\varphi_\sigma*\chi_E)|\,d\mathcal{L}^{n+1}\\ &\qquad\le c\Big(\int_U \zeta|D(\varphi_\sigma*\chi_E)|\,d\mathcal{L}^{n+1}\\ &\qquad\qquad + \int_U \varphi_\sigma*\chi_E|D\zeta|\,d\mathcal{L}^{n+1}\Big).\end{aligned}$$

Then by 2.5 of Ch. 2 it follows that

$$\Big(\int_E \zeta^{\frac{n+1}{n}}\,d\mathcal{L}^{n+1}\Big)^{\frac{n}{n+1}} \le c\Big(\int_U \zeta\,d\mu_E + \int_E |D\zeta|\,d\mathcal{L}^{n+1}\Big),$$

and replacing ζ by as sequence ζ_k as in (4), we get for a.e. $\rho\in(0,\rho_1)$

$$\big(\mathcal{L}^{n+1}(E\cap B_\rho(0))\big)^{\frac{n}{n+1}} \le c\Big(\mu_E(B_\rho(0)) + \frac{d}{d\rho}\mathcal{L}^{n+1}(E\cap B_\rho(0))\Big),$$

which by (6) gives

$$\big(\mathcal{L}^{n+1}(E\cap B_\rho(0))\big)^{\frac{n}{n+1}} \le c\frac{d}{d\rho}\mathcal{L}^{n+1}(E\cap B_\rho(0)) \quad \text{a.e. } \rho\in(0,\rho_1).$$

Integration (using the fact that $\mathcal{L}^{n+1}(E\cap B_\rho(0))$ is non-decreasing) then implies

$$\mathcal{L}^{n+1}(E\cap B_\rho(0)) \ge c\rho^{n+1} \tag{10}$$

for all sufficiently small ρ. Repeating the same argument with $U\setminus E$ in place of E, we also deduce

$$\mathcal{L}^{n+1}(B_\rho(0)\setminus E) \ge c\rho^{n+1} \tag{11}$$

for all sufficiently small ρ. (10) and (11) evidently tell us that $\lambda=0$ in (9).

Now given any sequence $\rho_k\downarrow 0$, the argument above guarantees we can select a subsequence $\rho_{k'}$ such that $\chi_{\rho_{k'}^{-1}E}\to\chi_{\{x\in\mathbb{R}^{n+1}:x^{n+1}<0\}}$ in $L^1_{\text{loc}}(\mathbb{R}^{n+1})$. Hence $\chi_{\rho^{-1}E}\to\chi_{\{x\in\mathbb{R}^{n+1}:x^{n+1}<0\}}$ and (2) of the theorem is established. Then by 4.1, (1) and (2) we have

$$\mu_{\rho^{-1}E}\to\mu_{\{x\in\mathbb{R}^{n+1}:x^{n+1}<0\}} = \mathcal{H}^n \llcorner \{x\in\mathbb{R}^{n+1} : x^{n+1}=0\} \text{ as } \rho\downarrow 0$$

and the proof is complete. □

Chapter 4

Theory of Rectifiable *n*-Varifolds

Let M be a countably n-rectifiable, $\mathcal{H}^n$-measurable subset of $\mathbb{R}^{n+\ell}$, and let θ be a positive locally $\mathcal{H}^n$-integrable function on M. Corresponding to such a pair (M, θ) we define the rectifiable n-varifold $\underline{v}(M, \theta)$ to be simply the equivalence class of all pairs $(\widetilde{M}, \widetilde{\theta})$, where $\widetilde{M}$ is countably n-rectifiable with $\mathcal{H}^n\big((M \setminus \widetilde{M}) \cup (\widetilde{M} \setminus M)\big) = 0$ and where $\widetilde{\theta} = \theta$ $\mathcal{H}^n$-a.e. on $M \cap \widetilde{M}$.[1] θ is called the multiplicity function of $\underline{v}(M, \theta)$. $\underline{v}(M, \theta)$ is called an integer multiplicity if this multiplicity function is integer-valued $\mathcal{H}^n$-a.e.

In this chapter and in Ch.5 we develop the theory of general n-rectifiable varifolds, particularly concentrating on *stationary* (see §2 below) rectifiable n-varifolds, which generalize the notion of classical minimal submanifolds of $\mathbb{R}^{n+\ell}$. The key section is §3, in which we obtain the monotonicity formula; much of the subsequent theory is based on this and closely related formulae.

[1] We shall see later that this is essentially equivalent to Allard's ([All72]) notion of n-dimensional rectifiable varifold.

1 Basic Definitions and Properties

Associated to a rectifiable n-varifold $V = \underline{v}(M, \theta)$ (as described above) there is a Radon measure μ (called the weight measure of V) defined by

1.1 $$\mu_V = \mathcal{H}^n \llcorner \theta,$$

where we adopt the convention that $\theta \equiv 0$ on $\mathbb{R}^{n+\ell} \setminus M$. Thus for an $\mathcal{H}^n$-measurable set A,

$$\mu_V(A) = \int_{A \cap M} \theta \, d\mathcal{H}^n,$$

the mass (or weight) of the varifold V, $\mathbb{M}(V)$, is defined by

1.2 $$\mathbb{M}(V) = \mu_V(\mathbb{R}^{n+\ell}) = \int_M \theta \, d\mathcal{H}^n.$$

Notice that by virtue of 1.9 of Ch. 3, an abstract Radon measure μ is actually μ_V for some rectifiable varifold V if and only if μ has an approximate tangent space P_x with multiplicity $\theta(x) \in (0, \infty)$ for μ-a.e. $x \in \mathbb{R}^{n+\ell}$. (See the statement of 1.9 of Ch. 3 for the terminology.) In this case $V = \underline{v}(M, \theta)$, where $M = \{x : \Theta^{*n}(\mu, x) > 0\}$.

1.3 Definition: Given a rectifiable n-varifold $V = \underline{v}(M, \theta)$, we define the tangent space $T_x V$ to be the approximate tangent space of M (as defined in the statement of 1.9 of Ch. 3) whenever this exists; notice that this is independent of the choice of representative (M, θ) for the equivalence class $\underline{v}(M, \theta)$.

We also define, for $V = \underline{v}(M, \theta)$,

1.4 $$\operatorname{spt} V = \operatorname{spt} \mu_V,$$

and for any $\mathcal{H}^n$-measurable subset $A \subset \mathbb{R}^{n+\ell}$, $v \llcorner A$ is the rectifiable n-varifold defined by

1.5 $$V \llcorner A = \underline{v}\big(M \cap A, \, \theta\big|(M \cap A)\big).$$

Given $V = \underline{v}(M, \theta)$ and a sequence $V_k = \underline{v}(M_k, \theta_k)$ of rectifiable n-varifolds, we say that $V_k \to V$ provided $\mu_{V_k} \to \mu_V$ in the usual sense of Radon measures. (Notice that this is *not* varifold convergence in the sense of Ch.8.)

Next we want to discuss the notion of mapping a rectifiable n-varifold relative to a Lipschitz map. Suppose $V = \underline{v}(M, \theta)$, $M \subset U$, U open in $\mathbb{R}^{n+\ell}$, W open in $\mathbb{R}^{n+\ell}$ and suppose $f : \operatorname{spt} V \cap U \to W$ is proper[2], Lipschitz and 1:1. Then we define the image varifold $f_\# V$ by

1.6 $$f_\# V = \underline{v}\big(f(M), \theta \circ f^{-1}\big).$$

[2] i.e. $f^{-1}(K) \cap \operatorname{spt} V$ is compact whenever K is a compact subset of W

Since K compact $\Rightarrow$ $f^{-1}(K)$ compact and hence $f(M) \cap K = f(M \cap f^{-1}(K))$, the area formula 4.13 of Ch. 2 gives

1.7 $$\int_{f(M) \cap K} \theta \circ f^{-1} \, d\mathcal{H}^n = \int_{M \cap f^{-1}(K)} J_f^M \, \theta \, d\mathcal{H}^n,$$

so in particular $\theta \circ f^{-1}$ is locally $\mathcal{H}^n$-integrable in W, and therefore 1.6 does indeed define a rectifiable n-varifold in W. More generally if f satisfies the conditions above, except that f is not necessarily 1:1, then we define $f_\# V$ by

$$f_\# V = \underline{v}\big(f(M), \widetilde{\theta}\big),$$

where $\widetilde{\theta}$ is defined on $f(M)$ by $\sum_{x \in f^{-1}(y) \cap M} \theta(x)$ $\big(= \int_{f^{-1}(y) \cap M} \theta \, d\mathcal{H}^0\big)$. Notice that $\widetilde{\theta}$ is locally $\mathcal{H}^n$-integrable in W by virtue of the area formula (see §2 of Ch. 3), and in fact

1.8 $$\mathbb{M}(f_\# V) = \int_{f(M)} \widetilde{\theta} \, d\mathcal{H}^n \equiv \int_M J_M f \, \theta \, d\mathcal{H}^n,$$

where $J_M f$ is the Jacobian of f relative to M as defined in §2 of Ch. 3. Thus, assuming $m \geq n$, we define

1.9 $$J_f^M(x) = \sqrt{\det \mathcal{J}(x)}.$$

where $\mathcal{J}(x)$ is the matrix with $D_{\tau_k} f(x) \cdot D_{\tau_\ell} f(x)$ in the k-th row and ℓ-th column ($\tau_1, \ldots, \tau_n$ any orthonormal basis for $T_x M$), $d^M f_x : T_x M \to \mathbb{R}^{n+\ell}$ is the linear map induced by f as described in §2 of Ch. 3, and $(d^M f_x)^* : \mathbb{R}^m \to T_x M$ denotes the adjoint of $d^M f_x$.

2 First Variation

Suppose $\{\phi_t\}_{-\varepsilon < t < \varepsilon}$ $(\varepsilon > 0)$ is a 1-parameter family of diffeomorphisms of an open set U of $\mathbb{R}^{n+\ell}$ satisfying

2.1 $$\begin{cases} \phi_0 = 1_U, \ \exists \text{ compact } K \subset U \text{ such that } \phi_t | U \setminus K = 1_{U \setminus K} \ \forall t \in (-\varepsilon, \varepsilon) \\ (x, t) \to \phi_t(x) \text{ is a smooth map} : U \times (-\varepsilon, \varepsilon) \to U. \end{cases}$$

Then if $V = \underline{v}(M, \theta)$ is a rectifiable n-varifold and if $K \subset U$ is compact as in 2.1 above, we have, according to 1.8,

$$\mathbb{M}\big(\phi_{t\#}(V \llcorner K)\big) = \int_{M \cap K} J_M \phi_t \, \theta \, d\mathcal{H}^n,$$

and we can compute the *first variation* $\frac{d}{dt}\underline{\mathrm{M}}(\phi_{t\#}(V \llcorner K))\Big|_{t=0}$ exactly as in §3 of Ch. 2. We thus deduce

2.2
$$\frac{d}{dt}\mathbb{M}(\phi_{t\#}(V \llcorner K))\Big|_{t=0} = \int_M \operatorname{div}_M X \, d\mu_V,$$

where $X_{|x} = \frac{\partial}{\partial t}\phi(t,x)\Big|_{t=0}$ is the initial velocity vector for the family $\{\phi_t\}$ and where $\operatorname{div}_M X$ is as in §4 of Ch. 2:

2.3
$$\operatorname{div}_M X = \textstyle\sum_{j=1}^{n+\ell} \nabla_j^M X^j \; (= \textstyle\sum_{j=1}^{n+\ell} e_j \cdot (\nabla^M X^j)).$$

($\nabla^M X^j$ as in §2 of Ch. 3)

We say that V is *stationary in U* if the first variation vanishes in U. That is, by 2.2, the definition is as follows:

2.4 Definition: $V = \underline{v}(M,\theta)$ is stationary in U if $\frac{d}{dt}\underline{\mathrm{M}}(\phi_{t\#}(V \llcorner K))\Big|_{t=0} = 0$ for every family $\{\varphi_t\}$ as in 2.1; of course by 2.2 this is equivalent to the requirement $\int \operatorname{div}_M X \, d\mu_V = 0$ for any C^1 vector field X on U having compact support in U.

More generally let N be an $(n+\ell)$-dimensional C^2 submanifold of $\mathbb{R}^{n+L}$ ($\ell \le L$), U an open subset of $\mathbb{R}^{n+L}$ such that $(\overline{N} \setminus N) \cap U = \emptyset$, $M \subset N$, and $\{\varphi_t\}_{-1 \le t \le 1}$ a 1-parameter family of diffeomorphisms $U \to U$ such that

2.5
$$\begin{cases} \varphi_t(x) = \varphi(t,x) \text{ is a } C^2 \text{ map of } (t,x) \in (-1,1) \times U \to U \\ \qquad\qquad \text{with } \varphi_t(U \cap N) \subset U \cap N \ \forall t \in (-1,1) \\ \varphi_0(x) = x \ \forall x \in U, \ \varphi_t(x) = x \ \forall (t,x) \in (-1,1) \times U \setminus K, \end{cases}$$

where K is a compact subset of U. Then we have the following definition:

2.6 Definition: V is stationary in $U \cap N$ if $\frac{d}{dt}\mathbb{M}(\phi_{t\#}(V \llcorner K))\Big|_{t=0} = 0$ whenever $\{\phi_t\}$ are as in 2.5.

As already mentioned in the discussion preceding 5.8 of Ch. 2, for each C^1 vector field X on U with $X_{|x} \in T_x N \ \forall x \in N \cap U$, there is a 1-parameter family φ_t as in 2.5 with initial velocity $\frac{\partial}{\partial t}\varphi_t(x)|_{t=0} = X_{|x}$ for each x in x. Thus in this context, when $V = \underline{v}(M,\theta)$ is stationary in $U \cap N$, we can compute, using the area formula exactly as in 2.2 above, that $V = \underline{v}(M,\theta)$ is stationary in $U \cap N$ if and only if

$$\int_M \operatorname{div}_M X \, d\mu_V = 0$$

for each C^1 vector field X on U with X tangent to N at each point of $N \cap U$; that is, $X \in C_c^1(U, \mathbb{R}^{n+L})$ with $X_{|x} \in T_x N \ \forall x \in N \cap U$. On the other hand, by exactly

the computation of 5.9 of Ch. 2 (which did not depend on smoothness of M), we can start with *any* C^1 vector field X on U and compute (as in 5.9 of Ch. 2) that

$$\operatorname{div}_M X = \operatorname{div}_M X^T - \sum_{i=1}^n \overline{H}^N \cdot X$$

at all points $x \in M$ where M has an approximate tangent space $T_x M$, where X^T is C^1 with compact support in U and tangent to N at each point of $N \cap U$ and, as in 5.9 of Ch. 2, $\overline{H}^N = \sum_{i=1}^n \overline{B}_x(\tau_i,\tau_i)$, with $\tau_1,\dots,\tau_n$ any orthonormal basis for $T_x M$. Thus in fact we conclude that V is stationary in $U \cap N \iff$

2.7
$$\int_M \operatorname{div}_M X \, d\mu_V = -\int_M X \cdot \overline{H}^N \, d\mu_V \text{ for each } X \in C_c^1(U, \mathbb{R}^{n+L}).$$

3 Monotonicity Formulae in the Stationary Case

In this section we assume that U is open in $\mathbb{R}^{n+\ell}$, $V = \underline{v}(M,\theta)$ is stationary in U, which means Definition 2.4 holds, i.e.

3.1
$$\int_M \operatorname{div}_M X \, d\mu_V = 0$$

whenever X is a C^1 vector field on U with compact support in U. We proceed to extract important information from this identity by taking specific choices of the vector function $X = (X_1, \dots, X_{n+\ell})$. In fact we begin by choosing $X_x = \gamma(r)(x-\xi)$, where $\xi \in U$ is fixed, $r = |x-\xi|$, and $\gamma : \mathbb{R} \to [0,1]$ is a $C^1(\mathbb{R})$ function with

$$\gamma'(t) \le 0 \ \forall t, \ \gamma(t) \equiv 1 \text{ for } t \le \rho, \ \gamma(t) \equiv 0 \text{ for } t \ge R,$$

where $R > \rho > 0$ and $B_R(\xi) \subset U$.

For any $f \in C^1(U)$ and any $x \in M$ such that $T_x M$ exists (see 1.6, 1.9 of Ch. 3) we have (by 2.1 of Ch. 3) $\nabla^M f(x) = \sum_{j,k=1}^{n+\ell} e^{jk} D_k lf(x) e_j$, where $D_k f$ denotes the partial derivative $\frac{\partial f}{\partial x^k}$ of f taken in U and where (e^{jk}) is the matrix of the orthogonal projection of $\mathbb{R}^{n+\ell}$ onto $T_x M$. Thus, writing $\nabla_j^M = e_j \cdot \nabla^M$ (as in §2), with the above choice of X we deduce

3.2
$$\operatorname{div}_M X = \textstyle\sum_{j=1}^{n+\ell} \nabla_j^M X^j = \gamma(r) \textstyle\sum_{j=1}^{n+\ell} e^{jj} + r\gamma'(r) \textstyle\sum_{j,k=1}^{n+\ell} e^{jk} \frac{(x^j - \xi^j)}{r} \frac{(x^k - \xi^k)}{r}.$$

Since $(e^{j\ell})$ represents orthogonal projection onto $T_x M$ we have $\sum_{j=1}^{n+\ell} e^{jj} = n$ and

$$\textstyle\sum_{j,k=1}^{n+\ell} e^{j\ell} \frac{(x^j - \xi^j)}{r} \frac{(x^k - \xi^k)}{r} = \left|p_{T_x V}\left(\frac{x-\xi}{r}\right)\right|^2 = 1 - \left|D^\perp r\right|^2,$$

where $D^\perp r$ denote the orthogonal projection of $r^{-1}(x-\xi) = Dr$ (which is a vector of length $= 1$) onto $(T_x V)^\perp$. Writing $\mu = \mu_V$, the formula 3.1 thus yields

3.3 $$n\int \gamma(r)\, d\mu + \int r\gamma'(r)\, d\mu = \int r\gamma'(r)\big|(Dr)^\perp\big|^2\, d\mu,$$

provided $B_R(\xi) \subset U$ and $\rho \in (0, R]$, which we subsequently assume. Now take $\varepsilon \in (0,1)$ and a C^1 function $\varphi : \mathbb{R} \to [0,1]$ such that $\varphi(t) \equiv 1$ for $t \le 1$, $\varphi(t) = 0$ for $t \ge 1+\varepsilon$ and $\varphi'(t) \le 0$ for all t. Then we can use 3.3 with $\gamma(r) = \varphi(r/\rho)$ and $\rho < R/(1+\varepsilon)$. Since

$$r\gamma'(r) = r\rho^{-1}\varphi'(r/\rho) = -\rho\tfrac{\partial}{\partial\rho}\big[\varphi(r/\rho)\big]$$

this gives

$$nI(\rho) - \rho I'(\rho) = -\rho J'(\rho),\ \rho < R/(1+\varepsilon),$$

where

$$I(\rho) = \int_M \varphi(r/\rho)\, d\mu,\ J(\rho) = \int_M \varphi(r/\rho)|D^\perp r|^2\, d\mu.$$

Thus, multiplying by ρ^{-n-1} and rearranging, we have

3.4 $$\tfrac{d}{d\rho}\big(\rho^{-n} I(\rho)\big) = \rho^{-n} J'(\rho).$$

Since $(1+\varepsilon)^{-n} r^{-n}\frac{\partial}{\partial\rho}\big[\varphi(r/\rho)\big] \le \rho^{-n}\frac{\partial}{\partial\rho}\big[\varphi(r/\rho)\big] \le r^{-n}\frac{\partial}{\partial\rho}\big[\varphi(r/\rho)\big]$, by integration in 3.4 we get

3.5 $$\begin{aligned}\int_M (1+\varepsilon)^{-n} r^{-n}(\varphi(r/\rho) - \varphi(r/\sigma))|D^\perp r|^2\, d\mu_V &\le \rho^{-n} I(\rho) - \sigma^{-n} I(\sigma) \\ &\le \int_M r^{-n}(\varphi(r/\rho) - \varphi(r/\sigma))|D^\perp r|^2\, d\mu_V\end{aligned}$$

Thus letting φ decrease to the indicator function of the interval $(-\infty, 1]$ by letting $\varepsilon \downarrow 0$, we obtain

3.6 $$\rho^{-n}\mu_V\big(B_\rho(\xi)\big) - \sigma^{-n}\mu_V\big(B_\sigma(\xi)\big) = \int_{B_\rho(\xi)\setminus B_\sigma(\xi)} \frac{\big|D^\perp r\big|^2}{r^n}\, d\mu_V,\ 0 < \sigma \le \rho < R,$$

provided $B_R(\xi) \subset U$. We also of course have the differentiated version of this, namely

3.7 $$\frac{d}{d\rho}\big(\rho^{-n}\mu_V\big(B_\rho(\xi)\big)\big) = \frac{d}{d\rho}\int_{B_\rho(\xi)} \frac{\big|D^\perp r\big|^2}{r^n}\, d\mu_V,$$

which holds in the distribution sense on $(0, R)$. Since both $\mu\big(B_\rho(\xi)\big)$ and $\int_{B_\rho(\xi)} \frac{|D^\perp r|^2}{r^n}$ are *increasing* in ρ, the differentiated version 3.7 also holds in the *classical sense* for a.e. $\rho \in (0, R)$, assuming $B_R(\xi) \subset U$.

3.6 is the fundamental monotonicity identity. In particular 3.6 tells us that the ratio

3.8 $$\rho^{-n}\mu\big(B_\rho(\xi)\big) \text{ is increasing in } \rho,\ 0 < \rho < R,$$

and hence the density

3.9 $$\Theta^n(\mu_V, \xi) = \lim_{\rho\downarrow 0} \frac{\mu_V(B_\rho(\xi))}{\omega_n \rho^n} \text{ exists and is real for } \textit{every}\ \xi \in U,$$

and by letting $\sigma \downarrow 0$ in 3.6 we also have

3.10 $$(\omega_n\rho^n)^{-1}\mu_V\big(B_\rho(\xi)\big) - \Theta^n(\mu_V, \xi) = \omega_n^{-1}\int_{B_\rho(\xi)} \frac{\big|D^\perp r\big|^2}{r^n}\, d\mu_V,\ 0 < \rho < R,$$

and in particular $\int_{B_\rho(\xi)} \frac{|D^\perp r|^2}{r^n}\, d\mu_V < \infty$. We also claim the upper semi-continuity

3.11 $$\Theta^n(\mu_V, \xi) \ge \limsup_{x\to\xi} \Theta^n(\mu_V, x), \quad \xi \in U.$$

To check this take and $\rho, \varepsilon > 0$ with $B_{\rho+\varepsilon}(\xi) \subset U$ and any sequence $\xi_j \to \xi$. Then $B_\rho(\xi_j) \subset B_{\rho+\varepsilon}(\xi)$ for all sufficiently large j, and hence using the monotonicity 3.8 we have

$$\Theta^n(\mu_V, \xi_j) \le (\omega_n\rho^n)^{-1}\mu_V(B_\rho(\xi_j)) \le (\omega_n\rho^n)^{-1}\mu_V(B_{\rho+\varepsilon}(\xi))$$

for all sufficiently large j, and hence

$$\limsup_{j\to\infty} \Theta^n(\mu_V, \xi_j) \le (\omega_n\rho^n)^{-1}\mu_V(B_{\rho+\varepsilon}(\xi)).$$

Letting $\varepsilon \downarrow 0$ we then have $\limsup_{j\to\infty}\Theta^n(\mu_V, \xi_j) \le (\omega_n\rho^n)^{-1}\mu_V(B_\rho(\xi))$, and finally, by letting $\rho \downarrow 0$, we obtain 3.11 as claimed.

Since $V = \underline{v}(M, \theta)$ and $\Theta^n(\mu_V, x) = \theta(x)$ for $\mathcal{H}^n$-a.e. $x \in M$ (by Remark 1.8(2)), 3.11 enables us to choose "canonical representatives" M_V, θ_V for V, so that $V = \underline{v}(M_V, \theta_V)$, where

3.12 $$M_V = \{x \in U : \Theta^n(\mu_V, x) > 0\} \text{ and } \theta_V(x) = \Theta^n(\mu_V, x)\ \forall x \in U.$$

Since θ_V is then upper semi-continuous in U by 3.11 we then have

3.13 $$\{x \in M_V : \theta_V(x) \ge \alpha\} \text{ is relatively closed in } U \text{ for each } \alpha > 0$$

and in particular M_V itself is relatively closed in U (and in fact equal to spt $V \cap U$) in case there exists $\alpha > 0$ with $\theta(x) \ge \alpha$ for $\mathcal{H}^n$-a.e. $x \in M$ (and then of course $\theta_V(x) \ge \alpha$ for *every* $x \in M_V$ by 3.11).

We now want to generalize this discussion to a context which includes of varifolds which are stationary in an $(n+\ell)$-dimensional C^2 submanifold N rather than in

$\mathbb{R}^{n+\ell}$, as discussed in §2 above. We in fact introduce the concept of *generalized mean curvature vector* for the varifold $V = \underline{v}(M, \theta)$ as follows:

3.14 Definition: Let $V = \underline{v}(M, \theta)$ be a rectifiable varifold in the open set $U \subset \mathbb{R}^{n+\ell}$. Then we say that V has generalized mean curvature vector $\underline{H}$ if

(‡) $$\int_M \operatorname{div}_M X \, d\mu_V = -\int_M X \cdot \underline{H} \, d\mu_V \ \forall X \in C_c^1(U, \mathbb{R}^{n+\ell}).$$

Thus V is stationary in U if and only if it has generalized mean curvature zero.

Notice also V is stationary in $U \cap N$, where U is open in $\mathbb{R}^{n+L}$ and N is a C^2 $(n+\ell)$-dimensional submanifold of $\mathbb{R}^{n+L}$, if and only if V has generalized mean curvature $\overline{H}^N$ in U, with $\overline{H}^N$ as in 2.7 of the previous section.

We want to show that the above monotonicity discussion generalizes to the case when $V = \underline{v}(M, \theta)$ has *bounded* generalized mean curvature $\underline{H}$. So suppose that there is constant Λ such that

3.15 $$|\underline{H}| \le \Lambda \text{ on } M \cap U.$$

We can then proceed on the left side of 3.14 (‡) exactly as in the case $\underline{H} = 0$ with the same choices of X, thus giving

3.16 $$\tfrac{d}{d\rho}(\rho^{-n} I(\underline{r})) = \rho^{-n} \tfrac{d}{d\rho} J(\rho) - E_0(\rho), \ 0 < \rho < R,$$

assuming $B_R(\xi) \subset U$, where I, J are as in 3.4 and $E_0(\rho) = \rho^{-n} \int_U \rho^{-1}(x - \xi) \cdot \underline{H} \varphi(r/\rho) \, d\mu_V$, so that, since $\varphi(r/\rho) = 0$ for $r > (1+\varepsilon)\rho$,

$$-\Lambda_\varepsilon \rho^{-n} I(\rho) \le E_0(\rho) \le \Lambda_\varepsilon \rho^{-n} I(\rho),$$

where $\Lambda_\varepsilon = (1+\varepsilon)\Lambda$, and hence $E_0(\rho) = E(\rho)\rho^{-n} I(\rho)$, where $E(\rho) \in [-\Lambda_\varepsilon, \Lambda_\varepsilon]$ for each $\rho \in (0, R)$. Thus, after multiplying each side of 3.16 by the integrating factor $F(\rho) = e^{\int_0^\rho E(t)\, dt} \in [e^{-\Lambda_\varepsilon \rho}, e^{\Lambda_\varepsilon \rho}]$, we obtain

$$e^{-\Lambda_\varepsilon R} \rho^{-n} \tfrac{d}{d\rho} J(\rho) \le \tfrac{d}{d\rho}(F(\rho)\rho^{-n} I(\rho)) \le e^{\Lambda_\varepsilon R} \rho^{-n} \tfrac{d}{d\rho} J(\rho).$$

Then taking φ as in 3.5 and integrating from σ to ρ as in the case $\underline{H} = 0$ and then letting $\varepsilon \downarrow 0$ as we did before, we obtain (analogous to 3.6)

$$F(\rho)\rho^{-n} \mu_V(B_\rho(\xi)) - F(\sigma)\sigma^{-n} \mu_V(B_\sigma(\xi)) = G(\sigma, \rho) \int_{B_\rho(\xi) \setminus B_\sigma(\xi)} r^{-n} |D^\perp r|^2 \, d\mu_V,$$

with $G(\sigma, \rho) \in [e^{-\Lambda R}, e^{\Lambda R}]$. Thus we have proved the following:

3.17 Theorem. *If U is open in $\mathbb{R}^{n+\ell}$, if $B_R(\xi) \subset U$ and V has generalized mean curvature vector $\underline{H}$ in U with $|\underline{H}| \le \Lambda$, then*

$$F(\rho)\rho^{-n} \mu_V(B_\rho(\xi)) - F(\sigma)\sigma^{-n} \mu_V(B_\sigma(\xi)) = G(\sigma, \rho) \int_{B_\rho(\xi) \setminus B_\sigma(\xi)} r^{-n} |D^\perp r|^2 \, d\mu_V,$$

for all $0 < \sigma < \rho < R$, where $F(\rho) \in [e^{-\Lambda\rho}, e^{\Lambda\rho}]$ for all $0 < \rho < R$ and $G(\sigma, \rho) \in [e^{-\Lambda R}, e^{\Lambda R}]$ for all $0 < \sigma \le \rho < R$.

Then notice that since $F(\rho) \in [e^{-\Lambda\rho}, e^{\Lambda\rho}]$, we again conclude that $\Theta^n(\mu_V, \xi)$ exists for all $\xi \in N \cap U$ and is an upper semi-continuous function on $N \cap U$ by very straightforward modifications to the previous argument for $\underline{H} = 0$.

We conclude this section with an observation that will be important in our later discussion of tangent cones: Namely, if $\delta > 0$ and $0 < \sigma < \rho$ are given, and if, instead of the boundedness condition for $\underline{H}$ assumed above, we merely require that

$$\underline{H} \in L^1_{\text{loc}}(U) \text{ and } \tau^{1-n} \int_{B_\tau(\xi)} |\underline{H}| \, d\mu_V \le \delta \text{ for all } \sigma \le \tau \le \rho,$$

then in place of the identity 3.7 we evidently have

$$\Big| \frac{d}{d\tau} \Big((\tau^{-n} \mu_V(B_\tau(\xi))) - \int_{B_\tau(\xi)} \frac{|D^\perp r|^2}{r^n} \, d\mu_V \Big) \Big| \le \delta \tau^{-1}, \quad \sigma < \tau < \rho,$$

assuming of course that $B_\rho(\xi) \subset U$. By integrating this gives

3.18 $$\Big| \Big(\rho^{-n} \mu_V(B_\rho(\xi)) - \sigma^{-n} \mu_V(B_\sigma(\xi))\Big) - \int_{B_\rho(\xi) \setminus B_\sigma(\xi)} r^{-n} |D^\perp r|^2 \, d\mu_V \Big| \le \delta \log(\rho/\sigma), \quad \sigma < \tau < \rho.$$

4 Monotonicity Formulae for L^p Mean Curvature

Here we continue to assume that $V = \underline{v}(M, \theta)$ is a rectifiable varifold in U (U open in $\mathbb{R}^{n+\ell}$) with generalized mean curvature vector $\underline{H}$ in U, but now we assume $\underline{H}$ is merely in L^p function rather than L^∞ as in the previous section.

Using the identity 3.14 (‡) we proceed exactly as in the previous section to obtain (Cf. 3.16)

4.1 $$\tfrac{d}{d\rho}(\rho^{-n} I(\rho)) = \rho^{-n} \tfrac{d}{d\rho} J(\rho) - \rho^{-n} \int_U \rho^{-1}(x - \xi) \cdot \underline{H} \, \varphi(r/\rho) \, d\mu_V, \ 0 < \rho < R,$$

assuming $B_R(\xi) \subset U$, where I, J are as in 3.4. But now we assume only an L^p condition with $p > n$ instead of a bound on H. Specifically we assume

4.2 $$p > n \text{ and } \Big(R^{p-n} \int_{B_R(\xi)} |H|^p \, d\mu_V\Big)^{1/p} \le \kappa\Lambda,$$

where Λ is a constant to be chosen and $\kappa = \frac{p}{4(p-n)}$ appears in front of Λ merely as a convenience to simplify the form of the main monotonicity identity below.

Observe that then by using Hölder's inequality we have (since $\varphi(r/\rho) = 0$ for $r < (1+\varepsilon)\rho$)

$$\begin{aligned}
&\big|\rho^{-n}\int_U \rho^{-1}(x-\xi)\cdot \underline{H}\,\varphi(r/\rho)\,d\mu_V\big| \\
&\qquad \le (1+\varepsilon)\rho^{-n}\|H\|_{L^p(\mu_V \llcorner B_R(\xi))}(I(\rho))^{1-1/p} \\
&\qquad = (1+\varepsilon)\rho^{-n/p}\|H\|_{L^p(\mu_V \llcorner B_R(\xi))}\big(\rho^{-n}I(\rho)\big)^{1-1/p} \\
&\qquad = (1+\varepsilon)R^{-1}(\rho/R)^{-n/p}\big(R^{p-n}\int_{B_R(\xi)}|H|^p\,d\mu_V\big)^{1/p}\big(\rho^{-n}I(\rho)\big)^{1-1/p} \\
&\qquad \le 2\kappa\Lambda R^{-1}(\rho/R)^{-n/p}\big(1+\rho^{-n}I(\rho)\big),\quad \rho < R/(1+\varepsilon),
\end{aligned}$$

where at the last step we used $a^{1-1/p} \le 1+a$, valid for $a \ge 0$. Thus 4.1 can be written

4.3 $$\tfrac{d}{d\rho}(\rho^{-n}I(\rho)) = \rho^{-n}\tfrac{d}{d\rho}J(\rho) - F_0(\rho)\big(1+\rho^{-n}I(\rho)\big),$$

where
$$|F_0(\rho)| \le 2\kappa\Lambda R^{-1}(\rho/R)^{-n/p},$$
so after multiplying through by the integrating factor $F(\rho) = e^{\int_0^\rho F_0(t)\,dt}$ in 4.3, we obtain

4.4 $$\tfrac{d}{d\rho}\big(F(\rho)\rho^{-n}I(\rho) + E(\rho)\big) = F(\rho)\rho^{-n}\tfrac{d}{d\rho}J(\rho),\quad 0<\rho<R/(1+\varepsilon),$$

where $E(\rho) = F(\rho) - F(0) = F(\rho) - 1$. Since $\int_0^\rho |F_0(t)|\,dt \le \frac{1}{2}\Lambda(\rho/R)^{1-n/p}$, and $|F(\rho)-F(0)| = |\int_0^\rho F'(t)\,dt| \le e^{\frac{1}{2}\Lambda}\int_0^\rho |F_0(t)|\,dt$, we then have the bounds

4.5 $$\begin{cases} e^{-\Lambda} \le e^{-\frac{1}{2}\Lambda(\rho/R)^{1-n/p}} \le F(\rho) \le e^{\frac{1}{2}\Lambda(\rho/R)^{1-n/p}} \le e^{\Lambda}, \\ |E(\rho)| \le \dfrac{1}{2}e^{\frac{1}{2}\Lambda}\Lambda(\rho/R)^{1-n/p},\quad 0<\rho\le R. \end{cases}$$

In particular if $\Lambda \le 1$ we have

4.6 $$|E(\rho)| \le \Lambda(\rho/R)^{1-n/p},\ 0<\rho\le R.$$

Thus we can proceed exactly as before, integrating from σ to ρ and letting $\varepsilon \downarrow 0$ in order to conclude

4.7 $$\begin{aligned}\Big(F(\rho)\rho^{-n}\mu_V(B_\rho(\xi)) + E(\rho)\Big) &- \Big(F(\sigma)\sigma^{-n}\mu_V(B_\sigma(\xi)) + E(\sigma)\Big) \\ &= G(\sigma,\rho)\int_{B_\rho(\xi)\setminus B_\sigma(\xi)} \tfrac{|D^\perp r|^2}{r^n}\,d\mu_V,\ G(\sigma,\rho)\in[e^{-\Lambda},e^{\Lambda}],\end{aligned}$$

for all $0<\sigma\le\rho\le R$, where 4.5 holds. In particular

4.8 $$F(\rho)\rho^{-n}\mu_V(B_\rho(\xi)) + E(\rho) \text{ is increasing in } \rho,\ 0<\rho\le R,$$

with E, F as in 4.5. 4.5 and 4.8 evidently enable us to argue precisely as in §3, to conclude that $\Theta^n(\mu_V,\xi) = \lim_{\rho\downarrow 0}(\omega_n\rho^n)^{-1}\mu_V(B_\rho(\xi))$ exists $\forall\,\xi\in U$ and

4.9 $$\Theta^n(\mu_V,\xi) \text{ is an upper semi-continuous function on } U.$$

4.10 Remarks: (1) In the case of $\underline{H}\in L^p_{\rm loc}(\mu_V)$ with $p>n$, if $\theta\ge 1$ μ_V-a.e. in U, then $\Theta^n(\mu_V,x)\ge 1$ at *each* point of $\operatorname{spt}\mu_V\cap U$, and hence we can write $V = \underline{v}(M_*,\theta_*)$ where $M_* = \operatorname{spt}\mu_V\cap U$, $\theta_*(x) = \Theta^n(\mu_V,x)$, $x\in U$. Thus V is represented in terms of a relatively *closed* countably n-rectifiable set in U with a multiplicity function which is *upper semi-continuous* in U.

(2) Notice that 4.7 gives bounds of the form $\mu(B_\sigma(\xi)) \le \beta\sigma^n$, $0<\sigma\le R$ for suitable constant β. Such an inequality implies

$$\int_{B_\rho(\xi)\setminus B_\sigma(\xi)} |x-\xi|^{-\alpha}\,d\mu \le \begin{cases} n\beta(n-\alpha)^{-1}(\rho^{n-\alpha}-\sigma^{n-\alpha}),\ 0<\alpha<n \\ \beta\log(\rho/\sigma),\ \alpha=n. \end{cases}$$

for any $0<\sigma<\rho<R$. This is proved for $0<\alpha<n$ by using the following general fact with $f(x) = |x-\xi|^{-1}$, $t_0 = \rho^{-1}$, and with $n-\alpha$ in place of α, and with

4.11 Lemma. *If X is an abstract space, μ is a measure on X with $\mu(X)<\infty$, $f\in L^1(\mu)$, $f\ge 0$, and if $A_t = \{x\in X : f(x)>t\}$, then*
$$\int_{t_0}^\infty t^{\alpha-1}\mu(A_t bigr)\,dt = \alpha^{-1}\int_{A_{t_0}}(f^\alpha - t_0^\alpha)\,d\mu,\quad 0<\alpha<n$$
$$\int_{t_0}^\infty t^{-1}\mu(A_t)\,dt = \int_{A_{t_0}}\log(f/t_0)\,d\mu$$
for each $t_0>0$.

Proof: Since $\int_{t_0}^\infty t^{\alpha-1}\mu(A_t)\,dt = \int_{t_0}^\infty\int_X t^{\alpha-1}\chi_{A_t}(x)\,d\mu(x)\,dt$, this is proved simply by applying Fubini's theorem on the product space $A_{t_0}\times[t_0,\infty)$. $\square$

5 Poincaré and Sobolev Inequalities

In this section[3] we continue to assume that $V = \underline{v}(M,\theta)$ has generalized mean curvature $\underline{H}$ in U, and we again write μ for μ_V. We shall also assume $\theta\ge 1$ μ-a.e. $x\in U$ (so that (by the comments in 4.10) $\Theta^n(\mu,x)\ge 1$ everywhere in $\operatorname{spt}\mu\cap U$ if $\underline{H}\in L^p_{\rm loc}(\mu)$ for some $p>n$).

[3]Note: The results of this section are not needed in the sequel

We begin by considering the possibility of repeating the argument of the previous section, but with $X_x = h(x)\gamma(r)(x-\xi)$ (rather than $X_x = \gamma(r)(x-\xi)$ as before), where h is a non-negative function in $C^1(U)$. In computing $\operatorname{div}_M X$ we will get the additional term $\gamma(r)(x-\xi)\cdot\nabla^M h$, and other terms will be as before with an additional factor $h(x)$ everywhere. Thus in place of 3.7 we get

5.1 $$\frac{\partial}{\partial\rho}\big(\rho^{-n}\widetilde{I}(\rho)\big) = \rho^{-n}\frac{\partial}{\partial\rho}\int \big|(Dr)^\perp\big|^2 h\,\varphi(r/\rho)\,d\mu + \rho^{-n-1}\int (x-\xi)\cdot\big[\nabla^M h + \underline{H}h\big]\,\varphi(r/\rho)\,d\mu$$

where now $\widetilde{I}(\rho) = \int \varphi(r/\rho)h\,d\mu$.

Thus

$$\frac{\partial}{\partial\rho}\big[\rho^{-n}\widetilde{I}(\rho)\big] \ge \rho^{-n-1}\int (x-\xi)\cdot\big(\nabla^M h + \underline{H}h\big)\,\varphi(r/\rho)\,d\mu \equiv R \text{ say.}$$

We can estimate the right-side R here in two ways: if $|\underline{H}| \le \Lambda$ we have

5.2 $$R \ge -\rho^{-n-1}\int r\big|\nabla^M h\big|\,\varphi(r/\rho)\,d\mu - (\Lambda\rho)\rho^{-n}\widetilde{I}(\rho).$$

Alternatively, without any assumption on $\underline{H}$ we can clearly estimate

5.3 $$R \ge -\rho^{-n-1}\int r\Big(\big|\nabla^M h\big| + h\big|\underline{H}\big|\Big)\,\varphi(r/\rho)\,d\mu.$$

If we use 5.2 in 5.1 and integrate (making use of 4.11) we obtain (after letting φ increase to the characteristic function of $(-\infty,1)$ as before)

5.4 $$\frac{1}{\omega_n\sigma^n}\int_{B_\sigma(\xi)} h\,d\mu \le e^{\Lambda\rho}\Big(\frac{1}{\omega_n\rho^n}\int_{B_\rho(\xi)} h\,d\mu + \frac{1}{n\omega_n}\int_{B_\rho(\xi)}\frac{\big|\nabla^M h\big|}{|x-\xi|^{n-1}}\,d\mu\Big),$$

provided $B_\rho(\xi) \subset U$ and $0<\sigma<\rho$.

If instead we use 5.3 then we similarly get

$$\frac{1}{\omega_n\sigma^n}\int_{B_\sigma(\xi)} h\,d\mu \le \frac{1}{\omega_n\rho^n}\int_{B_\rho(\xi)} h\,d\mu + \frac{1}{\omega_n}\int_\sigma^\rho \tau^{-n-1}\int_{B_\tau(\xi)} r\Big(\big|\nabla^M h\big| + h\big|\underline{H}\big|\Big)\,d\mu\,d\tau.$$

and hence (by 4.11 again)

5.5 $$\frac{1}{\omega_n\sigma^n}\int_{B_\sigma(\xi)} h\,d\mu \le \frac{1}{\omega_n\rho^n}\int_{B_\rho(\xi)} h\,d\mu + \frac{1}{n\omega_n}\int_{B_\rho(\xi)}\frac{\big(\big|\nabla^M h\big| + h\big|\underline{H}\big|\big)}{|x-\xi|^{n-1}}\,d\mu$$

provided $B_\rho(\xi) \subset U$ and $0<\sigma<\rho$.

If we let $\sigma \downarrow 0$ in 5.4 then we get (since $\Theta(\mu,\xi) \ge 1$ for $\xi \in \operatorname{spt}\mu$)

$$h(\xi) \le e^{\Lambda\rho}\Big(\frac{1}{\omega_n\rho^n}\int_{B_\rho(\xi)} h\,d\mu + \frac{1}{n\omega_n}\int_{B_\rho(\xi)}\frac{\big|\nabla^M h\big|}{|x-\xi|^{n-1}}\Big),\ \xi \in \operatorname{spt}\mu,\ B_\rho(\xi) \subset U.$$

We now state our Poincaré-type inequality.

5.6 Theorem. *Suppose $h \in C^1(U)$, $h \ge 0$, $B_{2\rho}(\xi) \subset U$, $|\underline{H}| \le \Lambda$, $\theta \ge 1$ μ-a.e. in U and*

$$\mu\big\{x \in B_\rho(\xi) : h(x) > 0\big\} \le (1-\alpha)\omega_n\rho^n,\ e^{\Lambda\rho} \le 1+\alpha$$

for some $\alpha \in (0,1)$. Suppose also that

(‡) $$\mu\big(B_{2\rho}(\xi)\big) \le \Gamma\rho^n,\ \Gamma > 0.$$

Then there are constants $\beta = \beta(n,\alpha,\Gamma) \in (0,1/2)$ and $c = c(n,\alpha,\Gamma) > 0$ such that

$$\int_{B_{\beta\rho}(\xi)} h\,d\mu \le c\rho\int_{B_\rho(\xi)}\big|\nabla^M h\big|\,d\mu.$$

Proof: To begin we take β to be an arbitrary parameter in $(0,1/2)$ and apply 5.5 with $\eta \in B_{\beta\rho}(\xi) \cap \operatorname{spt}\mu$ in place of ξ. This gives

(1)
$$h(\eta) \le e^{\Lambda(1-\beta)\rho}\Big(\frac{1}{\omega_n((1-\beta)\rho)^n}\int_{B_{(1-\beta)\rho}(\eta)} h\,d\mu + \frac{1}{n\omega_n}\int_{B_{(1-\beta)\rho}(\xi)}\frac{\big|\nabla^M h\big|}{|x-\eta|^{n-1}}\,d\mu\Big)$$
$$\le e^{\Lambda\rho}\Big(\frac{1}{\omega_n((1-\beta)\rho)^n}\int_{B_\rho(\xi)} h\,d\mu + \frac{1}{n\omega_n}\int_{B_\rho(\xi)}\frac{\big|\nabla^M h\big|}{|x-\eta|^{n-1}}\,d\mu\Big).$$

Now let γ be a fixed C^1 non-decreasing function on $\mathbb{R}$ with $\gamma(t)=0$ for $t \le 0$ and $\gamma(t) \le 1$ everywhere, and apply (1) with $\gamma(h-t)$ in place of h, where $t \ge 0$ is fixed. Then by (1)

$$\gamma(h(\eta)-t) \le \frac{1+\alpha}{n\omega_n}\int_{B_\rho(\xi)}\frac{\omega'(h-t)\big|\nabla^M h\big|}{|x-\eta|^{n-1}}\,d\mu + (1-\alpha^2)(1-\beta)^{-n}.$$

Selecting β small enough so that $(1-\beta)^{-n}\big(1-\alpha^2\big) \le 1-\alpha^2/2$, we thus get

(2) $$\frac{\alpha^2}{2} \le \frac{1+\alpha}{n\omega_n}\int_{B_\rho(\xi)}\frac{\gamma'(h-t)\big|\nabla^M h\big|}{|x-\eta|^{n-1}}\,d\mu$$

for any $\eta \in B_{\beta\rho}(\xi) \cap \operatorname{spt}\mu$ such that $\gamma(h(\eta)-t) \ge 1$. Now let $\varepsilon > 0$ and choose γ such that $\gamma(t) \equiv 1$ for $t \ge 1+\varepsilon$. Then (2) implies

$$1 \le c\int_{B_\rho(\xi)}\frac{\gamma'(h-t)\big|\nabla^M h\big|}{|x-\eta|^{n-1}}\,d\mu,\ \eta \in B_{\beta\rho}(\xi) \cap A_{t+\varepsilon},$$

where $A_\tau = \{y \in \operatorname{spt}\mu : h(y) > \tau\}$. Integrating over $A_{\tau+\varepsilon} \cap B_{\beta\rho}(\xi)$ we thus get (after interchanging the order of integration on the right)

$$\begin{aligned}(A_{t+\varepsilon} \cap B_{\beta\rho}(\xi)) &\le c \int_{B_\rho(\xi)} \gamma'(h(x)-t)\left|\nabla^M h(x)\right| \Big(\int_{B_{\beta\rho}(\xi)} \frac{1}{|x-\eta|^{n-1}}\, d\mu(\eta)\Big)\, d\mu(x)\\ &\le c\Gamma\rho \int_{B_\rho(\xi)} \gamma'(h-t)\left|\nabla^M h\right| d\mu\end{aligned}$$

by hypothesis 5.6(‡) and 4.10(2). Since $\gamma'(h(x)-t) = -\frac{\partial}{\partial t}\gamma(h(x)-t)$ we can now integrate over $t \in (0,\infty)$ to obtain (from 4.11) that

$$\int_{A_\varepsilon \cap B_{\beta\rho}(\xi)} (h-\varepsilon) \le c\Gamma\rho \int_{B_\rho(\xi)} \left|\nabla^M h\right| d\mu.$$

Letting $\varepsilon \downarrow 0$, we have the required inequality. □

Remark: If we drop the assumption that $\theta \ge 1$, then the above argument still yields

$$\int_{\{x:\theta(x)\ge 1\}\cap B_{\beta\rho}(\xi)} h\, d\mu \le c\rho \int_{B_\rho(\xi)} \left|\nabla^M h\right| d\mu.$$

We can also prove a Sobolev inequality as follows.

5.7 Theorem. *Suppose $h \in C_0^1(U)$, $h \ge 0$, and $\theta \ge 1$ μ-a.e. in U. Then*

$$(‡)\qquad \Big(\int h^{\frac{n}{n-1}}\, d\mu\Big)^{\frac{n-1}{n}} \le c \int \Big(\left|\nabla^M h\right| + h\left|\underline{H}\right|\Big)\, d\mu,\ c = c(n).$$

Note: c does not depend on k.

In the proof we shall need the following simple calculus lemma.

5.8 Lemma. *Suppose f, g are bounded and non-decreasing on $(0,\infty)$ and*

$$(‡)\qquad 1 \le \sigma^{-n} f(\sigma) \le \rho^{-n} f(\rho) + \int_0^\rho \tau^{-n} g(\tau)\, d\tau,\ 0 < \sigma < \rho < \infty.$$

then $\exists\rho$ with $0 < \rho < \rho_0 \equiv 2(f(\infty))^{1/n}$ ($f(\infty) = \lim_{\rho\uparrow\infty} f(\rho)$) such that

$$(‡‡)\qquad f(5\rho) \le \frac{1}{2} 5^n \rho_0\, g(\rho).$$

Proof of Lemma: Suppose (‡) is false for each $\rho \in (0,\rho_0)$. Then (‡‡) $\Rightarrow$

$$\begin{aligned}1 \le \sup_{0<\sigma<\rho_0} \sigma^{-n} f(\sigma) &\le \rho_0^{-n} f(\rho_0) + \frac{2.5^{-n}}{\rho_0} \int_0^{\rho_0} \rho^{-n} f(5\rho)\, d\rho\\ &\equiv \rho_0^{-n} f(\rho_0) + \frac{2}{5\rho_0} \int_0^{5\rho_0} \rho^{-n} f(\rho)\, d\rho\\ &\equiv \rho_0^{-n} f(\rho_0) + \frac{2}{5\rho_0} \Big(\int_0^{\rho_0} \rho^{-n} f(\rho)\, d\rho + \int_{\rho_0}^{5\rho_0} \rho^{-n} f(\rho)\, d\rho\Big)\\ &\le \rho_0^{-n} f(\infty) + \frac{2}{5} \sup_{0<\rho<\rho_0} \rho^{-n} f(\rho) + \frac{2}{5(n-1)} \rho_0^{-n} f(\infty).\end{aligned}$$

Thus

$$\frac{1}{2} \le \frac{1}{2} \sup_{0<\sigma<\rho_0} \sigma^{-n} f(\sigma) < 2\rho_0^{-n} f(\infty) = 2^{-n},$$

which is a contradiction. □

Continuation of the proof of 5.7: First note that because h has compact support in U, the formula 5.5 is actually valid here for all $0 < \sigma < \rho < \infty$. Hence we can apply the above lemma with the choices

$$\begin{aligned}f(\rho) &= \omega_n^{-1} \int_{B_\rho(\xi)} h\, d\mu,\\ g(\rho) &= \omega_n^{-1} \int_{B_\rho(\xi)} (\left|\nabla^M h\right| + h\left|\underline{H}\right|)\, d\mu,\end{aligned}$$

provided that $\xi \in \operatorname{spt}\mu$ and $h(\xi) \ge 1$.

Thus for each $\xi \in \{x \in \operatorname{spt}\mu : h(x) \ge 1\}$ we have $\rho < 2(\omega_n^{-1}\int_M h\, d\mu)^{1/n}$ such that

$$(1)\qquad \int_{B_{5\rho}(\xi)} h\, d\mu \le 5^n (\omega_n^{-1} \int_M h\, d\mu)^{1/n} \int_{B_\rho(\xi)} (\left|\nabla^M h\right| + h\left|\underline{H}\right|)\, d\mu.$$

Using the covering Lemma (3.4 of Ch. 1) we can select disjoint balls $B_{\rho_1}(\xi_1)$, $B_{\rho_2}(\xi_2)$, …, $\xi_1 \in \{\xi \in \operatorname{spt}\mu : h(\xi) \ge 1\}$ such that $\{\xi \in M : h(\xi) \ge 1\} \subset \cup_{j=1}^\infty B_{5\rho_j}(\xi_j)$. Then applying (1) and summing over j we have

$$\int_{\{x\in\operatorname{spt}\mu : h(x)\ge 1\}} h\, d\mu \le 5^n \Big(\omega_n^{-1} \int_M h\, d\mu\Big)^{1/n} \int_M (\left|\nabla^M h\right| + h\left|\underline{H}\right|)\, d\mu.$$

Next let γ be a non-decreasing $C^1(\mathbb{R})$ function such that $\gamma(t) \equiv 1$ for $t > \varepsilon$ and $\gamma(t) \equiv 0$ for $t < 0$, and use this with $\gamma(h-t)$, $t \ge 0$, in place of h. This gives

$$\mu(M_{t+\varepsilon}) \le 5^n \omega_n (\mu(M_t))^{1/n} \int_M (\gamma'(h-t)\left|\nabla^M h\right| + \gamma(h-t)\left|\underline{H}\right|)\, d\mu,$$

where

$$M_\alpha = \{x \in M : h(x) > \alpha\},\ \alpha \ge 0.$$

Multiplying this inequality by $(t+\varepsilon)^{\frac{1}{n-1}}$ and using the trivial inequality $(t+\varepsilon)^{\frac{1}{n-1}}\mu(M_t) \le \int_{M_t}(h+\varepsilon)^{\frac{1}{n-1}}\,d\mu$ on the right, we then get

$$(t+\varepsilon)^{\frac{1}{n-1}}\mu(M_{t+\varepsilon}) \le$$
$$5^n\omega_n^{-1/n}\Big(\int_M (h+\varepsilon)^{\frac{n}{n-1}}\,d\mu\Big)^{\frac{1}{n}}\Big(-\frac{d}{dt}\int_M \gamma(\xi-t)|\nabla^M h| + \int_{M_t}|\underline{H}|\,d\mu\Big).$$

Now integrate of $t \in (0,\infty)$ and use 4.11. This then gives

$$\int_{M_\varepsilon}\big(h^{\frac{n}{n-1}} - \varepsilon^{\frac{n}{n-1}}\big)\,d\mu \le 5^{n+1}\omega_n^{-1/n}\Big(\int_M (h+\varepsilon)^{\frac{n}{n-1}}\Big)^{\frac{1}{n}}\int_M \big(|\nabla^M h| + h|\underline{H}|\big)\,d\mu.$$

The theorem (with $c = 5^{n+1}\omega_n^{-1/n}$) now follows by letting $\varepsilon \downarrow 0$. □

5.9 Remark: Note that the inequality of 5.7 is valid without any boundedness hypothesis on $\underline{H}$: it suffices that $\underline{H}$ is merely in $L^1_{\rm loc}(\mu)$.

6 Miscellaneous Additional Consequences

Here $V = \underline{v}(M,\theta)$ is a rectifiable n-varifold in $\mathbb{R}^{n+\ell}$ with generalized mean curvature $\underline{H}$ in U, $U \subset \mathbb{R}^{n+\ell}$ open, as in Definition 3.14 of the present chapter. We first derive a preliminary property for V in case $\underline{H}$ is bounded.

6.1 Lemma. *Suppose $U = \mathbb{R}^{n+\ell}\setminus B_R(\xi)$ and $V \llcorner U$ has $L^1_{\rm loc}(\mu_V)$ generalized mean curvature $\underline{H}$ in U with $n^{-1}|\underline{H}(x)\cdot(x-\xi)| < 1$ μ_V-a.e. in U, and suppose also that* spt V *is compact. Then*

$$\text{spt}\, V \subset B_R(\xi).$$

(i.e. $V \llcorner U = 0$.)

Proof: Since spt V is compact it is easily checked that the identity (see §3)

$$n\int \gamma(r)\,d\mu_V + \int r\gamma'(r)\big(1-|D^\perp r|^2\big)\,d\mu_V = -\int \underline{H}(x)\cdot(x-\xi)\gamma(r)\,d\mu_V(x)$$

(where $r = |x-\xi|$) actually holds for any non-negative increasing $C^1(\mathbb{R})$ function γ with $\gamma(t) = 0$ for $t \le R+\varepsilon$. ($\varepsilon > 0$ arbitrary.) We see this as in §3, by substituting $X(x) = \psi(x)\gamma(r)(x-\xi)$, where $\psi \in C^1_c(\mathbb{R}^{n+\ell})$ with $\psi \equiv 1$ in a neighborhood of spt V. Since $1-|D^\perp r|^2 \ge 0$ and $|\underline{H}\cdot(x-\xi)| < n$ μ_V-a.e., we thus deduce $\int \gamma(r)\,d\mu_V = 0$ for any such γ. Since we may select γ so that $\gamma(t) > 0$ for $t > R+\varepsilon$, we thus conclude spt V ($\equiv$ spt μ_V) $\subset B_{R+\varepsilon}(\xi)$. Because $\varepsilon > 0$ was arbitrary, this proves the lemma.

6.2 Theorem (Convex hull property for stationary varifolds.) *Suppose* spt V *is compact and V is stationary in $\mathbb{R}^{n+\ell}\setminus K$, K compact. Then*

$$\text{spt}\, V \subset \textit{convex hull of } K.$$

Proof: The convex hull of K can be written as the intersection of all balls $B_R(\xi)$ with $K \subset B_R(\xi)$. Hence the result follows immediately from 6.1. □

The observation of the following lemma is important.

XXX Hausdorff distance sense convergence.

6.3 Lemma. *Suppose $\theta \ge 1$ μ-a.e. in U, $\underline{H} \in L^p_{\rm loc}(\mu)$ in U for some $p > n$. If the approximate tangent space T_xV (see §1) exists at a given point $x \in U$, then T_xV is a "classical" tangent plane for* spt μ *in the sense that*

$$\lim_{\rho\downarrow 0}\big(\sup\{\rho^{-1}\,\text{dist}\,(y, T_xV) : y \in \text{spt}\,\mu \cap B_\rho(x)\}\big) = 0.$$

Proof: For sufficiently small ρ_0 (where $0 < \rho_0 < R$ with $B_R(x) \subset U$), 4.7 (with $\sigma \downarrow 0$) and 4.10(1) evidently imply

(1) $$\omega_n^{-1}\rho^{-n}\mu(B_\rho(\xi)) \ge 1/2,\ 0 < \rho < \rho_0,\ \xi \in \text{spt}\,\mu_V.$$

Using this we are going to prove that if $\alpha \in (0,1/2)$ and $\rho \in (0,R)$ then

(2) $$\mu\big(B_\rho(x)\setminus\{y : \text{dist}(y,T_xV) < \varepsilon\rho\}\big) < \frac{\omega_n}{2}\alpha^n\rho^n$$
$$\Rightarrow \text{spt}\,\mu \cap B_{\rho/2}(x) \subset \{y : \text{dist}(y,T_xV) < (\varepsilon+\alpha)\rho\}.$$

Indeed if $\xi \in \{y : \text{dist}(y,T_xV) \ge (\varepsilon+\alpha)\rho\} \cap B_{\rho/2}(x)$, then $B_{\alpha\rho}(\xi) \subset B_\rho(x)\setminus\{y : \text{dist}(y,T_xV) < \varepsilon\rho\}$ and hence the hypothesis of (2) implies $\mu(B_{\alpha\rho}(\xi)) < \frac{1}{2}\omega_n\alpha^n\rho^n$. On the other hand (1) implies $\mu(B_{\alpha\rho}(\xi)) \ge \frac{1}{2}\omega_n\alpha^n\rho^n$, so we have a contradiction. Thus (2) is proved, and (2) evidently leads immediately to the required result. □

Chapter 5

The Allard Regularity Theorem

Here we discuss Allard's ([All72]) regularity theorem, which says roughly that if the generalized mean curvature of a rectifiable n-varifold $V = \underline{v}(M, \theta)$ is in $L^p_{\mathrm{loc}}(\mu_V)$ in U, $p > n$, if $\theta \ge 1$ μ_V-a.e. in u, if $\xi \in \operatorname{spt} V \cap U$, and if $\omega_n^{-1}\rho^{-n}\mu_V(B_\rho(\xi))$ is sufficiently close to 1 for *some* sufficiently small[1] ρ, then V is regular near ξ in the sense that spt V is a $C^{1,1-n/p}$ n-dimensional submanifold near ξ.

A key idea of the proof is to show that V is well-approximated by the graph of a harmonic function near ξ. We begin in the first section with a motivating discussion, where we consider smooth minimal surfaces with small C^1 norm, and discuss the fact that in such a classical setting harmonic functions do indeed give a very good approximation.

The rest of the chapter to devoted to Allard's theorem, beginning in §2 with a discussion of the fact that a stationary n-dimensional rectifiable varifold V in a ball $B_R(\xi) \subset \mathbb{R}^{n+\ell}$ which has mass density ratio $(\omega_n R^n)^{-1}\mu_V(B_R(\xi))$ close to 1 has

[1] Depending on $\|\underline{H}\|_{L^p(\mu_V)}$

nice affine approximation properties near every point in the support, and can be very well approximated by a Lipschitz graph with small Lipschitz constant. We in fact do this under the assumption that the generalized mean curvature has small L^p norm with $p > n$.

In §4 we show that the harmonic approximation lemma of §3 can be applied to the Lipschitz approximation of §2, leading to the "tilt-excess decay" theorem, which is the main step in the proof of the Allard theorem.

The idea of approximating by harmonic functions (in roughly the sense used here) goes back to De Giorgi [DG61] who proved a special case of the above theorem (when $k = 1$ and when V corresponds to the reduced boundary of a set of least perimeter—see the previous discussion in §4 of Ch. 3 and the discussion in §5 below). Almgren used analogous approximations in his work [Alm68] for arbitrary $k \geq 1$. Reifenberg [Rei60, Rei64] used approximation by harmonic functions in a rather different way in his work on regularity of minimal surfaces.

1 Harmonic Approximation in the Smooth Case

Suppose M is an n-dimensional C^2 submanifold of $\mathbb{R}^{n+\ell}$. We say that M is a minimal submanifold if its mean curvature vector $\underline{H}$ is identically zero. From the discussion in Ch.2 we have seen that this is exactly equivalent to the volume $\mathcal{H}^n(M)$ being stationary with respect to compactly supported perturbations of the identity. Thus, in the notation of §5 of Ch.2, M is minimal if and only if $\frac{d}{dt}\mathcal{H}^n(\varphi_t(M))|_{t=0} = 0$. We showed that this in turn is equivalent to the first variation identity $\int_M \operatorname{div}_M X\, d\mathcal{H}^n = 0$. In the present smooth case we can use the local graphical representations discussed in 4.4. Thus, modulo an orthogonal transformation of $\mathbb{R}^{n+\ell}$ we can locally write M as a graph of a C^2 function with values in $\mathbb{R}^\ell$ over a domain in $\mathbb{R}^n$. Thus for each $\xi \in M$ we can assume there are open sets $W \subset \mathbb{R}^{n+\ell}$ and a ball $B_\rho(\eta) \subset \mathbb{R}^n$ and a C^2 map $u : B_\rho(\eta) \to \mathbb{R}^\ell$ such that $u(\eta) = \xi$, $Du(\eta) = 0$ and $\operatorname{graph} u (= \{(x, u(x)) : x \in \breve{B}_\rho(\eta)\}) = M \cap W$. Then stationarity of M implies in particular that the area functional $\mathcal{A}(u) = \int_{B_\rho(\eta)} J_u\, d\mathcal{H}^n$ must be stationary with respect to compactly supported perturbations of u in $B_\rho(\eta)$, where J_u is the Jacobian of the graph map $x \in B_\rho(\eta) \mapsto (x, u(x)) \in \operatorname{graph} u = M \cap W$. Thus $J_u = \sqrt{\det \mathcal{J}}$, where $\mathcal{J}(x)$ is the $n \times n$ matrix $(D_i(x, u(x)) \cdot D_j(x, u(x)))$; i.e., the $n \times n$ matrix with entry $(e_i, D_i u(x)) \cdot (e_j, D_j u(x)) = \delta_{ij} + D_i u(x) \cdot D_j u(x)$ in the i-th row and j-th column. Thus

$$\mathcal{A}(u) = \int_{B_\rho(\eta)} \sqrt{\det(\delta_{ij} + D_i u \cdot D_j u)}\, d\mathcal{L}^n$$

and for $|Du| < \varepsilon_0$ (for suitably small $\varepsilon_0 = \varepsilon_0(n, \ell) \in (0, 1)$) we can use a Taylor series expansion to give

1.1 $$\mathcal{A}(u) = \int_{B_\rho(\eta)} (1 + \frac{1}{2}|Du|^2 + F(Du))\, d\mathcal{L}^n,$$

where $F = F(P)$ is a real analytic map of $n \times \ell$ matrices $P = (p_{ij})$ with $|P| < \varepsilon_0$ to $\mathbb{R}$ such that

1.2 $$|F(P)| \leq C|P|^4, \ |D_{p_{ij}} F(P)| \leq C|P|^3, \quad |P| \leq 1.$$

where C is a fixed constant depending only on n, ℓ.

Since $\mathcal{A}(u)$ is stationary with respect to compactly supported perturbations of u we have

$$\tfrac{d}{dt}\mathcal{A}(u + t\zeta)|_{t=0} = 0, \quad \zeta = (\zeta_1, \ldots, \zeta_\ell) \in C^1_0(B_\rho(\eta), \mathbb{R}^\ell),$$

where $C^1_0(B_\rho(\eta), \mathbb{R}^\ell)$ denotes the C^1 maps $\zeta : B_\rho(\eta) \to \mathbb{R}^\ell$ with $\zeta = 0$ on $\partial B_\rho(\eta)$. In view of 1.1, if $|Du| < \varepsilon_0$ this takes the form

$$\int_{B_\rho(\eta)} (1 + |Du|^2 + F(Du))^{-1/2} \textstyle\sum_{i=1}^n D_i u \cdot D_i \zeta\, d\mathcal{L}^n = \int_{B_\rho(\eta)} \sum_{i=1}^n \sum_{j=1}^\ell A_{ij}(Du) D_i \zeta_j),$$

for all $\zeta \in C^1_0(B_\rho(\eta), \mathbb{R})$, where $A_{ij}(P) = D_{p_{ij}} F(P)$. This can be rewritten

1.3 $$\int_{B_\rho(\eta)} \textstyle\sum_{i=1}^n D_i u \cdot D_i \zeta\, d\mathcal{L}^n = \int_{B_\rho(\eta)} \sum_{i=1}^n \sum_{j=1}^\ell \widetilde{A}_{ij}(Du) D_i \zeta_j),$$

where $\widetilde{A}_{ij}(P) = A_{ij}(P) - (1 - (1 + |P|^2 + F(P))^{-1/2}) P_{ij}$, so $|\widetilde{A}_{ij}(P)| \leq C|P|^3$. Integrating by parts, we get

$$\Delta u = \textstyle\sum_{i,j=1}^n a_{ij}(Du) D_i D_j u, \quad a_{ij}(Du) = O(|Du|^2).$$

It is therefore reasonable, so long as $|Du|$ is small, to expect that u is well approximated by a harmonic function. Indeed let us check this rigorously: Assume $|Du| < \varepsilon_0$ (ε_0 as above), and let v be the harmonic function on the ball $B_\rho(\eta)$ with $v = u$ on $\partial B_\rho(\eta)$—it is standard that such a harmonic function v exists and it is C^1 on $B_\rho(\eta)$ and $C^\infty(\breve{B}_\rho(\eta))$. Multiplying the equation $\Delta v = 0$ by ζ and integrating by parts over the ball $B_\rho(\eta)$, we obtain

1.4 $$\int_{B_\rho(\eta)} \textstyle\sum_{i=1}^n D_i v \cdot D_i \zeta\, d\mathcal{L}^n = 0, \quad \zeta \in C^1_0(B_\rho(\eta)).$$

Taking the difference between 1.3 and 1.4, we see then that

$$\int_{B_\rho(\eta)} \textstyle\sum_{i=1}^n D_i (u-v) \cdot D_i \zeta\, d\mathcal{L}^n = \int_{B_\rho(\eta)} \sum_{i=1}^n \sum_{j=1}^\ell A_{ij}(Du) D_i \zeta_j, \ \zeta \in C^1_0(B_\rho(\eta), \mathbb{R})$$

In this identity we take $\zeta = u - v$, so that

$$\int_{B_\rho(\eta)} |D(u-v)|^2\, d\mathcal{L}^n = \int_{B_\rho(\eta)} \sum_{i=1}^n \sum_{j=1}^\ell \widetilde{A}_{ij}(Du) D_i(u_j - v_j),$$

and using the Cauchy-Schwarz inequality $ab \le \frac{1}{2}a^2 + \frac{1}{2}b^2$ on the right side we get finally

$$\int_{B_\rho(\eta)} |D(u-v)|^2\, d\mathcal{L}^n \le \frac{1}{2}\int_{B_\rho(\eta)} \sum_{i,j}(A_{ij}(Du))^2 + \frac{1}{2}\int |D(u-v)|^2,$$

so

$$\int_{B_\rho(\eta)} |D(u-v)|^2\, d\mathcal{L}^n \le \int_{B_\rho(\eta)} \sum_{i,j}(A_{ij}(Du))^2.$$

That is, since $|\sum_{ij}(A_{ij}(P))^2| \le C|P|^6$ for $|P| < \varepsilon_0$, we obtain

1.5
$$\int_{B_\rho(\eta)} |D(u-v)|^2\, d\mathcal{L}^n \le C\int_{B_\rho(\eta)} |Du|^6.$$

This shows that indeed v is a very good approximation of u for $|Du|$ small: For example if $\sup_{B_\rho(\eta)} |Du| = \varepsilon < \varepsilon_0$, then 1.5 shows

$$(\omega_n \rho^n)^{-1}\int_{B_\rho(\eta)} |D(u-v)|^2\, d\mathcal{L}^n \le C\varepsilon^4 \int_{B_\rho(\eta)} |Du|^2\, d\mathcal{L}^n,$$

where (as in 1.2) C is a fixed constant depending only on n, ℓ, so that $\int_{B_\rho(\eta)} |D(u-v)|^2$ is much smaller than $\int_{B_\rho(\eta)} |Du|^2$ for ε small. Thus v is a very good approximation to u if ε is small.

So there is good motivation to think that harmonic approximation could be relevant in the study of the regularity of stationary rectifiable varifolds; indeed, as mentioned in the introduction to this chapter, we will show that such approximations are appropriate even in the more general context of rectifiable varifolds with generalized mean curvature in L^p, $p > n$.

2 Preliminaries, Lipschitz Approximation

In this section U is an open subset of $\mathbb{R}^{n+\ell}$, $B_R(0) \subset U$, $V = \underline{v}(M,\theta)$ is a rectifiable n-varifold with generalized mean curvature $\underline{H}$ in U (as in Definition 3.14 of Ch. 4).

We being with the following lemma relating tilt-excess and height; note that we do not need $\theta \ge 1$ for this.

2.1 Lemma. *Suppose $B_\rho(\xi) \subset U$. Then for any n-dimensional subspace $T \subset \mathbb{R}^{n+\ell}$ we have*

$$E(\xi,\rho/2,T) \le C\rho^{-n}\int_{B_\rho(\xi)} \Big(\frac{\operatorname{dist}(x-\xi,T)}{\rho}\Big)^2 d\mu + C\rho^{2-n}\int_{B_\rho(\xi)} |\underline{H}|^2\, d\mu,$$

where $C = C(n)$.

2.2 Remark: Note that in case $\rho^{-n}\mu(B_\rho(\xi)) \le \beta$, we can use the Hölder inequality to estimate the term $\int_{B_\rho(\xi)} |\underline{H}|^2\, d\mu$, giving

$$\rho^{2-n}\int_{B_\rho(\xi)} |\underline{H}|^2\, d\mu \le C\Big(\rho^{p-n}\int_{B_\rho(\xi)} |\underline{H}|^p\, d\mu\Big)^{1/p}, \quad p > 2,\ C = C(n,p,\beta).$$

Thus 2.1 gives

$$E(\xi,\rho/2,T) \le C\rho^{-n}\int_{B_\rho(\xi)} \Big(\frac{\operatorname{dist}(x-\xi,T)}{\rho}\Big)^2 d\mu + C\Big(\rho^{p-n}\int_{B_\rho(\xi)} |\underline{H}|^p\, d\mu\Big)^{2/p}$$

for $p \ge 2$, $C = C(n,p,\beta)$.

Proof of 2.1: It evidently suffices to prove the result with $\xi = 0$ and $T = \mathbb{R}^n \times \{0\}$. The proof simply involves making a suitable choice of X in the formula of 3.14 of Ch. 4. In fact we take

$$X = \zeta^2(x)x',\ x' = (0, x^{n+1}, \ldots, x^{n+\ell})$$

for $x = (x^1, \ldots, x^{n+\ell}) \in U$, where $\zeta \in C_c^1(U)$ with $\zeta \ge 0$ will be chosen below. By the definition of div_M (see §2 of Ch. 3) we have

$$\operatorname{div}_M x' = \sum_{i=n+1}^{n+\ell} e^{ii},\ \mu\text{-a.e. } x \in M,$$

where (e^{ij}) is the matrix of the projection p_{T_xM} (relative to the standard orthonormal basis for $\mathbb{R}^{n+\ell}$). Thus by the definition 3.14 of Ch. 4 of $\underline{H}$ we have

(1)
$$\int \sigma\zeta^2\, d\mu = \int \Big(-2\zeta \sum_{i=n+1}^{n+\ell}\sum_{j=1}^{n+\ell} x^i e^{ij} D_j\zeta - \zeta^2 x'\cdot \underline{H}\Big)\, d\mu,$$

with

(2)
$$\sigma \equiv \sum_{i=n+1}^{n+\ell} e^{ii} = \frac{1}{2}\sum_{i,j=1}^{n+\ell}(e^{ij} - \varepsilon^{ij})^2 = \frac{1}{2}|p_{T_xM} - p_{\mathbb{R}^n}|^2,$$

where (ε^{ij}) = matrix of $p_{\mathbb{R}^n}$ and where we used $(e^{ij})^2 = (e^{ij})$ and $\operatorname{trace}(e^{ij}) = n$. Also observe that $\varepsilon^{ij} = 0$ if $i > n$, so (1) can be written

(3)
$$\int \sigma\zeta^2\, d\mu = \int \Big(-2\zeta \sum_{i=n+1}^{n+\ell}\sum_{j=1}^{n+\ell} x^i (e^{ij} - \varepsilon^{ij}) D_j\zeta - \zeta^2 x'\cdot \underline{H}\Big)\, d\mu,$$

so

$$\int \sigma\zeta^2\,d\mu \le \int \big(2\sqrt{\sigma}|x'||\nabla\zeta|\zeta + |x'||\underline{H}|\zeta^2\big)\,d\mu.$$

Hence (using $ab \le \frac{1}{2}a^2 + \frac{1}{2}b^2$)

$$\int \sigma\zeta^2\,d\mu \le 4\int \big(|x'|^2|\nabla\zeta|^2 + |x'||\underline{H}|\zeta^2\big)\,d\mu,$$

The lemma now follows by choosing $\zeta \equiv 1$ in $B_{\rho/2}(0)$, $\zeta \equiv 0$ outside $B_\rho(0)$ and $|\nabla\zeta| \le 3/\rho$, and then noting that $|x'||\underline{H}| = (\rho^{-1}|x'|)(|\underline{H}|\rho) \le \frac{1}{2}\rho^{-2}|x'|^2 + \frac{1}{2}\big(|\underline{H}|\rho\big)^2$. $\square$

For the remainder of this section we continue to assume that $V = \underline{v}(M,\theta)$ has generalized mean curvature $\underline{H}$ and now we additionally assume, with $\delta \in (0,\frac{1}{2})$ a constant to be specified below and $\mu = \mu_V = \mathcal{H}^n \llcorner \theta$, the following:

2.3
$$\begin{cases} 1 \le \theta\ \mu\text{-a.e.},\ 0 \in \operatorname{spt} V,\ B_\rho(0) \subset U \\ \omega_n^{-1}\rho^{-n}\mu(B_\rho(0)) \le 1+\delta,\ \Big(\rho^{p-n}\int_{B_\rho(0)} |\underline{H}|^p\,d\mu\Big)^{1/p} \le \delta. \end{cases}$$

Notice that by the monotonicity formula 4.7 and the semi-continuity 4.9 of Ch. 4 we have, subject to 2.3, that $y \in \operatorname{spt} V \cap B_\delta(0) \Rightarrow 1 - C\delta \le (\omega_n\sigma^n)^{-1}\mu(B_\sigma(y)) \le (1+C\delta)(\omega_n(1-\delta)^n\rho^n)^{-1}\mu(B_{(1-\delta)\rho}(y)) \le 1 + C\delta$, because $B_{(1-\delta)\rho}(y) \subset B_\rho(0)$. Thus, for $\delta \le \delta_0(n,\ell,p) \in (0,\frac{1}{4}]$,

2.4 $$\tfrac{1}{2} \le 1 - C\delta \le \frac{\mu(B_\sigma(y))}{\omega_n\sigma^n} \le 1 + C\delta \le 2,\ 0 < \sigma \le (1-\delta)\rho,\ y \in \operatorname{spt} V \cap B_{\delta\rho}(0),$$

where $C = C(n,\ell,p)$.

Subject to conditions 2.3, we first establish a lemma which guarantees local affine approximations of the support at all points of spt V in the ball $B_{\delta\rho}(0)$ and in at radii $\sigma \le 2\delta\rho$:

2.5 Lemma (Affine Approximation Lemma.) *Suppose $\delta \in (0,\frac{1}{4}]$ and 2.3 holds. Then for each $\xi \in \operatorname{spt} V \cap B_{\delta\rho}(0)$, $\sigma \in (0, 2\delta\rho]$ there is an n-dimensional subspace $T = T(\xi,\sigma)$ with*

$$(\ddagger)\quad C^{-1}\Big(\sigma^{-n}\int_{B_{\sigma/2}(\xi)} |p_{T_xM} - p_T|^2\,d\mu_V(x)\Big)^{1/2} \le \sigma^{-1}\sup\{\operatorname{dist}(x, \xi + T) : x \in \operatorname{spt} V \cap B_\sigma(\xi)\} \le C\delta^{1/(2n+2)},$$

where $C = C(n,\ell,p)$.

Proof: Take any fixed $\sigma \in (0, 2\delta\rho]$ and $\xi \in \operatorname{spt} V \cap B_{\delta\rho}(0)$, and suppose for convenience of notation (by changing scale and translating the origin) that $\sigma = 1$ and

$\xi = 0$, so now, since $\delta \le \frac{1}{4}$ and hence $(1-\delta)\rho/(\delta\rho) \ge 3$, 2.4 holds for $\sigma \le 2$ and any $y \in \operatorname{spt} V \cap B_1(0)$. By 4.7 of Ch. 4 together with 2.4 above,

$$(1)\qquad \int_{B_2(y)} \big|p_{(T_xM)^\perp}(x-y)\big|^2 \le \int_{B_2(y)} \big|p_{(T_xM)^\perp}(x-y)\big|^2 |x-y|^{-n-2}\,d\mu \le C\delta$$

for $y \in \operatorname{spt} V \cap B_1(0)$. Next take $\alpha \in (0,1)$ (to be chosen shortly, but for the moment arbitrary). Recall the general principle that if K is compact and $\eta > 0$ then any maximal pairwise disjoint collection $B_{\eta/2}(y_j)$ with $y_j \in K$ will automatically have the property that $K \subset \cup_j B_\eta(y_j)$. Using this with $\eta = \delta^\alpha$ we have pairwise disjoint balls $B_{\delta^\alpha/2}(y_1), \ldots, B_{\delta^\alpha/2}(y_N)$ with $y_j \in \operatorname{spt} V \cap B_1(0)$ such that

$$(2)\qquad \operatorname{spt} V \cap B_1(0) \subset \cup_{j=1}^N B_{\delta^\alpha}(y_j).$$

Notice that then by 2.4 (with $\sigma = \delta^\alpha\rho$) we have

$$(3)\qquad C^{-1}\delta^{\alpha n} \le \mu(B_{\delta^\alpha/2}(y_j)) \le C\delta^{\alpha n},\ j = 1, \ldots, N,$$

and hence

$$C^{-1}N\delta^{\alpha n} \le \textstyle\sum_{j=1}^N \mu(B_{\delta^\alpha/2}(y_j)) = \mu(\cup_{j=1}^N B_{\delta^\alpha/2}(y_j)) \le C\mu(B_2(0)) \le 2C.$$

Thus $N \le C\delta^{-\alpha n}$, and so, by using (1) with $y = y_j$ and noting that $B_2(y_j) \supset B_1(0)$ for each j, we have

$$\int_{B_1(0)} \textstyle\sum_{j=1}^N \big|p_{T_xM^\perp}(x-y_j)\big|^2\,d\mu \le CN\delta = C\delta^{1-\alpha n}.$$

Thus for any given $k \ge 1$ we have

$$(4)\qquad \textstyle\sum_{j=1}^N \big|p_{T_xM^\perp}(x-y_j)\big|^2 \le Ck\delta^{1-\alpha n},$$

except possibly for a set of $x \in B_1(0) \cap \operatorname{spt} V$ of μ-measure $\le 1/k$. Since $\mu(B_{\delta^\alpha}(0)) \ge C^{-1}\delta^{\alpha n}$ by 2.4, we can select $k = C\delta^{-\alpha n}$, thus ensuring that (4) holds for some $x_0 \in \operatorname{spt} V \cap B_{\delta^\alpha}(0)$. So we have shown there is $x_0 \in \operatorname{spt} V \cap B_{\delta^\alpha}(0)$ with

$$(5)\qquad \textstyle\sum_{j=1}^N \big|p_{T_{x_0}M^\perp}(x_0-y_j)\big|^2 \le C\delta^{1-2\alpha n},$$

and hence

$$\big|p_{T_{x_0}M^\perp}(y_j - x_0)\big| \le C\delta^{\frac{1}{2}-\alpha n},\quad j = 1, \ldots, N.$$

Since $|x_0| < \delta^\alpha$, we then have

$$(6)\qquad \big|p_{T_{x_0}M^\perp}y_j\big| \le C(\delta^{\frac{1}{2}-\alpha n} + \delta^\alpha),\quad j = 1, \ldots, N.$$

Then, selecting α such that $\frac{1}{2} - \alpha n = \alpha$ (i.e. $\alpha = 1/(2n+2)$), we have shown that all the points $y_1, \ldots, y_N$ are in the $C\delta^{1/(2n+2)}$ neighborhood of the subspace $T_0 = T_{x_0}M$, and hence by (2) we have

$$\operatorname{dist}(y, T_0) \le C\delta^{1/(2n+2)} \quad \forall y \in \operatorname{spt} V \cap B_1(0).$$

and hence the second inequality in (‡) is proved with $T = T_0 = T_{x_0}M$.

The first inequality of (‡) then follows directly from Remark 2.2. □

We have the following important corollary of the above lemma:

2.6 Lemma (Lipschitz Approximation Lemma.) *Let $L \in (0,1]$ be given. There is $\beta = \beta(n,\ell,p) \in (0,\frac{1}{4}]$ such that if $\delta \in (0,(\beta L)^{2n+2}]$ and if* 2.3 *holds, if the subspaces $T(\sigma,\xi)$ are as in* Lemma 2.5*, if $\sigma_0 = \delta\rho$ and if we assume (without loss of generality) that $T(2\sigma_0, 0) = \mathbb{R}^n \times \{0\}$, then there is a Lipschitz $f : B^n_{\sigma_0/2}(0) \to \mathbb{R}^\ell$ with* $\operatorname{Lip} f \le L$, $\sup|f| \le C\delta^{1/(2n+2)}$ *and with*

$$\mu_V\big(B_{\sigma_0/2}(0) \cap (\operatorname{spt}\mu_V \setminus \operatorname{graph} f)\big) + \mathcal{H}^n\big(B_{\sigma_0/2}(0) \cap (\operatorname{graph} f \setminus \operatorname{spt}\mu_V)\big)$$
$$\le CL^{-2}\int_{B_{\sigma_0}(0)} \big|p_{T_xM} - p_{\mathbb{R}^n\times\{0\}}\big|^2\, d\mu_V, \quad C = C(n,\ell,p).$$

Proof: Assume 2.3 holds, where for the moment $\delta \in (0,\frac{1}{4}]$ is arbitrary and, using the notation of Lemma 2.5, let $T_0 = T(2\sigma_0, 0) = \mathbb{R}^n \times \{0\}$, where $\sigma_0 = \delta\rho$. Then by Lemma 2.5

$$(1) \qquad \sigma_0^{-n}\int_{B_{\sigma_0}(0)} |p_{T_xM} - p_{T_0}|^2\, d\mu_V(x) \le C\delta^{1/(n+1)},$$

with $C = C(n,\ell,p)$. Let

$$G = \{y \in \operatorname{spt}\mu_V \cap B_{3\sigma_0/4}(0) : \sup_{\sigma\in(0,\sigma_0/2]} \sigma^{-n}\int_{B_{\sigma/2}(y)} |p_{T_xM} - p_{T_0}|^2\, d\mu_V \le \beta^2L^2\},$$

where $\beta \in (0,\frac{1}{4}]$ is for the moment arbitrary but which we will choose shortly to depend only on n,ℓ,p. Thus $y \in \operatorname{spt}\mu_V \cap B_{3\sigma_0/4}(0) \setminus G \Rightarrow \exists\, \sigma \in (0,\sigma_0/2]$ with

$$(2) \qquad \beta^2L^2\sigma^n \le \int_{B_{\sigma/2}(y)} |p_{T_xM} - p_{T_0}|^2\, d\mu_V.$$

By the five-times covering lemma we can pick pairwise disjoint balls $B_{\sigma_j}(y_j)$ such that (2) holds with $\sigma = \sigma_j \in (0,\sigma_0/2]$ and $y = y_j \in \operatorname{spt}\mu_V \cap B_{\sigma_0}(0) \setminus G$, and such that

$$\operatorname{spt}\mu_V \cap B_{3\sigma_0/4}(0) \setminus G \subset \cup_j B_{5\sigma_j}(y_j).$$

Thus using (2) with $\sigma = \sigma_j$, $y = y_j$ and summing over j we obtain

$$\begin{aligned}\beta^2L^2\mu_V(B_{3\sigma_0/4}(0) \setminus G) &\le \beta^2L^2 \textstyle\sum_j \mu_V(B_{5\sigma_j}(y_j)) \\ &\le C\beta^2L^2 \textstyle\sum_j \sigma_j^n \le C\displaystyle\int_{\cup_j B_{\sigma_j/2}(y_j)} |p_{T_xM} - p_{T_0}|^2\, d\mu_V \\ &\le C\int_{B_{\sigma_0}(0)} |p_{T_xM} - p_{T_0}|^2\, d\mu_V.\end{aligned}$$

Thus

$$\mu_V(B_{3\sigma_0/4}(0) \setminus G) \le C\beta^{-2}L^{-2}\int_{B_{\sigma_0}(0)} |p_{T_xM} - p_{T_0}|^2\, d\mu_V.$$

We now claim that G is contained in the graph of a Lipschitz function. To check this, let y_1, y_2 be distinct points of G, let $\sigma = |y_1 - y_2|(\le \sigma_0)$ and observe that by definition of G and $T(y_1,\sigma)$ we have the two inequalities

$$\sigma^{-n}\int_{B_{\sigma/2}(y_1)} |p_{T_xM} - p_{T_0}|^2 \le \beta^2L^2, \quad \sigma^{-n}\int_{B_{\sigma/2}(y_1)} |p_{T_xM} - p_{T(y_1,\sigma)}|^2 \le \delta^{1/(n+1)}$$

and hence, since $|p_{T_0} - p_{T(y_1,\sigma)}|^2 \le 2|p_{T_xM} - p_{T_0}|^2 + 2|p_{T_xM} - p_{T(y_1,\sigma)}|^2$,

$$(3) \qquad |p_{T_0} - p_{T(y_1,\sigma)}| \le C(\beta L + \delta^{1/(2n+2)}).$$

Now by Lemma 2.5 we have

$$|p_{T(y_1,\sigma)}(y_1 - y_2)| = \operatorname{dist}(y_2, y_1 + T(y_1,\sigma)) \le C\delta^{1/(2n+2)}\sigma$$

and hence by (3)

$$\begin{aligned}(4) \qquad \operatorname{dist}(y_2, y_1 + T_0) &= |p_{T_0}(y_1 - y_2)| \\ &= |(p_{T(y_1,\sigma)} + (p_{T_0} - p_{T(y_1,\sigma)}))(y_1 - y_2)| \\ &\le \operatorname{dist}(y_2, y_1 + T(y_1,\sigma)) + |p_{T_0} - p_{T(y_1,\sigma)}|\sigma \\ &\le C(\beta L + \delta^{1/(2n+2)})\sigma \le C\beta L,\end{aligned}$$

assuming we take $\delta \le (\beta L)^{2n+2}$. This says that $|Q(y_2 - y_1)| \le C\beta L|y_1 - y_2| \le C\beta l(|Q(y_1 - y_2)| + |P(y_1 - y_2)|)$, where P, Q is denote the projections $y = (y^1, \ldots, y^{n+\ell})$ onto its first n and last ℓ coordinates respectively. Assuming $C\beta \le \frac{1}{2}$ we thus have $|Q(y_1) - Q(y_2)| \le 2C\beta L|P(y_1) - P(y_2)|$. In view of the arbitrariness of $y_1, y_2 \in G$ this says that G is contained in the graph of a Lipschitz function with Lipschitz constant $\le L$, provided we eventually choose $\beta = \beta(n,\ell,p)$ to satisfy the above restriction $C\beta \le \frac{1}{2}$; we also needed the restriction $\delta \le (\beta L)^{2n+2}$, so we need 2.3 to hold with $\delta \in (0,(\beta L)^{2n+2}]$.

Now by the Lipschitz Extension Theorem 1.2 of Ch.2, we thus have

$$(5) \qquad G \subset \operatorname{graph} f, \text{ where } f : \mathbb{R}^n \to \mathbb{R}^\ell \text{ with } \operatorname{Lip} f \le C\beta L\,(\le L),$$

and

$$\mu_V\big(B_{3\sigma_0/4}(0)\setminus G\big)\le CL^{-2}\int_{B_{\sigma_0}(0)}\big|p_{T_xM}-p_{T_0}\big|^2\,d\mu_V. \tag{6}$$

Of course, since $T_0=T_0(2\sigma_0,0)$, we have by Lemma 2.5 that $|f^j(x)|\le C\delta^{1/(2n+2)}$ for all x such that $(x,f(x))\in G$, so, by replacing f^j by

$$\widetilde{f}^j=\max\{\min\{f^j,C\delta^{1/(2n+2)}\},-C\delta^{1/(2n+2)}\},$$

we can assume

$$\sup|f|\le C\delta^{1/(2n+2)}. \tag{7}$$

It thus remains only to prove

$$\mathcal{H}^n\big((F\setminus\operatorname{spt}\mu)\cap B_{\sigma_0/2}(0)\big)\le CL^{-2}\int_{B_{\sigma_0}(0)}|p_{T_xM}-p_{T_0}|^2\,d\mu_V, \tag{8}$$

where $F=\operatorname{graph} f$ with f as in (5),(6),(7). To check this, take any $\eta\in(F\setminus\operatorname{spt}\mu)\cap B_{\sigma_0/2}(0)$ and let $\sigma\in(0,\sigma_0/4]$ be such that $B_\sigma(\eta)\cap\operatorname{spt}\mu=\emptyset$ and $B_{2\sigma}(\eta)\cap\operatorname{spt}\mu\ne\emptyset$. (Such σ exists because $0\in\operatorname{spt}\mu$, and $\eta\in B_{\sigma_0/2}(0)$.) Then the monotonicity identity 4.7 of Ch. 4 implies

$$\begin{aligned}\mu(B_{3\sigma}(\eta))&=\mu(B_{3\sigma}(\eta))-\mu(B_\sigma(\eta))\\&\le C\sigma^n\int_{B_{3\sigma}(\eta)\setminus B_\sigma(\eta)}|x-\eta|^{-n}\Big|p_{(T_xM)^\perp}\Big(\frac{x-\eta}{|x-\eta|}\Big)\Big|^2\,d\mu+C\delta\sigma^n.\end{aligned} \tag{9}$$

Now since $\operatorname{spt}\mu\cap B_{2\sigma}(\eta)\ne\emptyset$, 2.3 implies $\mu(B_{3\sigma}(\eta))\ge\frac12\omega_n\sigma^n$, and hence (9) gives, for suitable $\delta=\delta(p,n,\ell)$,

$$\begin{aligned}\sigma^n&\le C\int_{B_\sigma(\eta)}\Big|p_{(T_xM)^\perp}\Big(\frac{x-\eta}{\sigma}\Big)\Big|^2\,d\mu\\&\le C\Big(\int_{B_\sigma(\eta)}\Big|p_{(\mathbb{R}^n\times\{0\})^\perp}\Big(\frac{x-\eta}{\sigma}\Big)\Big|^2\,d\mu+\int_{B_\sigma(\eta)}|p_{T_xM}-p_{\mathbb{R}^n\times\{0\}}|^2\,d\mu\Big)\\&\le C\Big(\int_{B_\sigma(\eta)\cap F}\Big|p_{(\mathbb{R}^n\times\{0\})^\perp}\Big(\frac{x-\eta}{\sigma}\Big)\Big|^2\,d\mu+\mu(B_\sigma(\eta)\setminus F)\\&\qquad+\int_{B_\sigma(\eta)}|p_{T_xM}-p_{\mathbb{R}^n\times\{0\}}|^2\,d\mu\Big),\end{aligned}$$

where we use $p_{T^\perp}(x)=x-p_T(x)$ for any subspace $T\subset\mathbb{R}^{n+\ell}$. Since

$$\big|p_{(\mathbb{R}^n\times\{0\})^\perp}\big(\frac{x-y}{\sigma}\big)\big|\le C\beta \text{ for } x,y\in F\cap B_\sigma(\eta),$$

(because $\operatorname{Lip} f\le\beta L$), and $\mu(B_\sigma(\eta))\le 2\omega_n\sigma^n$ by 2.4, this implies

$$\sigma^n\le C\big(\beta L\sigma^n+\mu(B_\sigma(\eta)\setminus F)+\int_{B_\sigma(\eta)}|p_{T_xM}-p_{\mathbb{R}^n\times\{0\}}|^2\,d\mu\big).$$

With β chosen appropriately (depending only on n,ℓ,p), we can arrange that $C\beta\le$

$\frac12$, and hence this gives

$$\sigma^n\le C\big(\mu(B_\sigma(\eta)\setminus F)+\int_{B_\sigma(\eta)}\big|p_{T_xM}-p_{\mathbb{R}^n}\times\{0\}\big|^2\,d\mu\big). \tag{10}$$

Now observe that the collection of such balls $B_\sigma(\eta)$ by definition cover all of $B_{\sigma_0/2}(0)\cap F\setminus\operatorname{spt}\mu$, so by the 5-times covering lemma we can find a pairwise disjoint collection $B_{\sigma_j}(\eta_j)$ with

$$\sigma_j^n\le C\big(\mu(B_{\sigma_j}(\eta_j)\setminus F)+\int_{B_{\sigma_j}(\eta_j)}|p_{T_xM}-p_{\mathbb{R}^n}|^2\,d\mu\big). \tag{11}$$

for each j and $B_{\sigma_0/2}(0)\cap F\setminus\operatorname{spt}\mu\subset\cup_jB_{5\sigma_j}(\eta_j)$. Since F is the graph of the Lipschitz function f with $\operatorname{Lip} f\le1$, we of course have $\mathcal{H}^n(B_{5\sigma_j}(\eta_j)\cap F)\le C\sigma_j^n$ for each j, hence by (11)

$$\begin{aligned}\mathcal{H}^n(B_{\sigma_0/2}(0)\cap F\setminus\operatorname{spt}\mu)&\le\mathcal{H}^n(F\cap(\cup_jB_{5\sigma_j}(\eta_j)))\le\textstyle\sum_j\mathcal{H}^n(F\cap B_{5\sigma_j}(\eta_j))\\&\le C\textstyle\sum_j\big(\mu(B_{\sigma_j}(\eta_j)\setminus F)+\int_{B_{\sigma_j}(\eta_j)}|p_{T_xM}-p_{\mathbb{R}^n}|^2\,d\mu\big)\\&\le C\big(\mu(\cup_jB_{\sigma_j}(\eta_j)\setminus F)+\int_{\cup_jB_{\sigma_j}(\eta_j)}|p_{T_xM}-p_{\mathbb{R}^n}|^2\,d\mu\big)\\&\le C\big(\mu(B_{3\sigma_0/4}(0)\setminus F)+\int_{B_{\sigma_0}(0)}|p_{T_xM}-p_{\mathbb{R}^n}|^2\,d\mu\big)\\&\le C\int_{B_{\sigma_0}(0)}|p_{T_xM}-p_{\mathbb{R}^n}|^2\,d\mu\end{aligned}$$

by (6), so (8) is established. □

2.7 Corollary. *If the notation and assumptions are as in* Lemma 2.6, *then the parameter* $\beta=\beta(n,\ell,p)$ *can be chosen so that*

$$\sup_{\xi\in\operatorname{spt}\mu\cap B_{\sigma_0/2}(0),\sigma\in(0,\sigma_0/2]}\sigma^{-n}\int_{B_\sigma(\xi)}\big|p_{T_xM}-p_{\mathbb{R}^n\times\{0\}}\big|^2\,d\mu\le\beta^2L^2$$
$$\Rightarrow\operatorname{spt}\mu\cap B_{\sigma_0/4}(0)=\operatorname{graph} f\cap B_{\sigma_0/4}(0)$$

for some Lipschitz map $f:B^n_{\sigma_0/4}(0)\to\mathbb{R}^\ell$ *with* $\operatorname{Lip} f\le L$, $\sup|f|\le C\delta^{1/(2n+2)}\sigma_0$.

Proof: Note that if $\sup_{\xi\in B_{\sigma_0/2}(0),\sigma\le\sigma_0/2}\int_{B_\sigma(\xi)}\big|p_{T_xM}-p_{\mathbb{R}^n\times\{0\}}\big|^2\,d\mu\le\beta^2L^2$ then the set G in the above proof of Lemma 2.6 by definition includes all of $\operatorname{spt}\mu\cap B_{\sigma_0/2}(0)$, so $\operatorname{spt}\mu\cap B_{\sigma_0/2}(0)\subset\operatorname{graph} f$ with $\operatorname{Lip} f\le L$ and $\sup|f|\le C\delta^{1/(2n+2)}\sigma_0$. Further, if $\eta\in\operatorname{graph} f\cap B_{\sigma_0/4}(0)\setminus\operatorname{spt}\mu$ then inequality (10) in the above proof gives

$$\sigma^n\le C\int_{B_\sigma(\eta)}|p_{T_xM}-p_{\mathbb{R}^n\times\{0\}}|^2\,d\mu\le C\beta^2L^2\sigma^n\le C\beta^2\sigma^n$$

for some $\sigma\in(0,\sigma_0/4]$. This is evidently impossible with $\beta=\beta(n,\ell,p)$ chosen so that $C\beta^2\le\frac12$, so with such β we have $\operatorname{graph} f\cap B_{\sigma_0/4}(0)\setminus\operatorname{spt}\mu=\emptyset$ □

3 Approximation by Harmonic Functions

The main result we shall need is given in the following lemma, which is an almost trivial consequence of Rellich's theorem:

3.1 Lemma. *Given any $\varepsilon > 0$ there is a constant $\delta = \delta(n, \varepsilon) > 0$ such that if $f \in W^{1,2}(B)$, $B \equiv \breve{B}_1(0) =$ open unit ball in $\mathbb{R}^n$, satisfies*

$$\int_B |\nabla f|^2 \le 1, \quad \left|\int_B \nabla f \cdot \nabla \zeta \, d\mathcal{L}^n\right| \le \delta \sup |\nabla \zeta|$$

for every $\zeta \in C_c^\infty(B)$, then there is a harmonic function u on B such that $\int_B |\nabla u|^2 \le 1$ and

$$\int_B (u - f)^2 \le \varepsilon.$$

Proof: Suppose the lemma is false. Then we can find $\varepsilon > 0$ and a sequence $\{f_k\} \in W^{1,2}(B)$ such that

$$\left|\int_B \nabla f_k \cdot \nabla \zeta \, d\mathcal{L}^n\right| \le k^{-1} \sup |\nabla \zeta| \tag{1}$$

for each $\zeta \in C_c^\infty(B)$, and

$$\int_B \left|\nabla f_k\right|^2 \le 1,$$

but so that

$$\int_B \left|f_k - u\right|^2 > \varepsilon \tag{2}$$

whenever u is a harmonic function on B with $\int_B |\nabla u|^2 \le 1$. Let $\lambda_k = \omega_n^{-1} \int_B f_k \, d\mathcal{L}^n$. Then by the Poincaré inequality (see e.g. [GT01]) we have

$$\int_B \left|f_k - \lambda_k\right|^2 \le C \int_B \left|\nabla f_k\right|^2 \le C,$$

and hence, by Rellich's theorem (see [GT01]), we have subsequence $\{k'\} \subset \{k\}$ such that $f_{k'} - \lambda_{k'} \to w$ with respect to the $\mathcal{L}^2(B)$ norm and $\nabla f_{k'} \rightharpoonup \nabla w$ weakly in L^2, where $w \in W^{1,2}(B)$ with $\int_B |\nabla w|^2 \le 1$. By the weak convergence of $\nabla f_{k'}$ to ∇w and by (1) we evidently have

$$\int_B \nabla w \cdot \nabla \zeta \, d\mathcal{L}^n = \lim \int_B \nabla f_k \cdot \nabla \zeta \, d\mathcal{L}^n = 0$$

for each $\zeta \in C_c^\infty(B)$. Thus w is harmonic in B and $\int_B \left|f_{k'} - w - \lambda_{k'}\right|^2 \to 0$. Since $w + \lambda_{k'}$ is harmonic, this contradicts (2). $\square$

We also recall the following standard estimates for harmonic functions (which follow directly from the mean-value property—see e.g. [GT01]): If u is harmonic on $B \equiv B_\sigma(0)$, then

$$\sup_{B_{\sigma/2}(0)} \sigma^q \left|D^q u\right| \le C \sigma^{-n/2} \|u\|_{L^2(B)} \tag{3.2}$$

for each integer $q \ge 0$, where $C = C(q, n)$. Indeed applying this with Du in place of u we get

$$\sup_{B_{\sigma/2}(0)} \sigma^{q-1} \left|D^q u\right| \le C \|\nabla u\|_{L^2(B)} \tag{3.3}$$

for $q \ge 1$. Using 3.2, 3.3 and an order 2 Taylor series expansion for u, we see that if ℓ is the affine approximation to u given by $\ell(x) = u(0) + x \cdot \nabla u(0)$ then

$$\begin{cases} |\ell(0)| = |u(0)| \le C\sigma^{-n/2} \|u\|_{L^2(B)}, \ |\nabla \ell| = |\nabla u(0)| \le C\sigma^{-n/2} \|\nabla u\|_{L^2(B)} \\ \sup_{B_{\eta\sigma}(0)} |u - \ell| \le (\eta\sigma)^2 \sup_{B_{\eta\sigma}} |D^2 u| \le (\eta\sigma)^2 | \sup_{B_{\sigma/2}} |D^2 u| \le C \eta^2 \sigma^{1-n/2} \|\nabla u\|_{L^2(B)} \end{cases} \tag{3.4}$$

for $\eta \in (0, \frac{1}{4}]$, where $C = C(n)$ is independent of η.

4 The Tilt-Excess Decay Lemma

In this section we continue to assume V has generalized mean curvature $\underline{H}$ in U (as in Definition 3.14 of Ch. 4), and we write μ for μ_V.

If $B_\sigma(\xi) \subset U$ we define the tilt-excess $E(\xi, \sigma, T)$ (relative to the rectifiable n-varifold $V = \underline{v}(M, \theta)$) by

$$E(\xi, \sigma, T) = \sigma^{-n} \int_{B_\sigma(\xi)} \left|p_{T_x M} - p_T\right|^2 d\mu_V, \text{[2]} \tag{4.1}$$

T is an n-dimensional subspace of $\mathbb{R}^{n+\ell}$. Thus E measures the mean-square deviation of the approximate tangent space $T_x M$ away from the given subspace T. Notice that if we have $T = \mathbb{R}^n$ then, in terms of the $(n + \ell) \times (n + \ell)$ orthogonal projection matrices (e^{ij}) and (ε^{ij}) for $T_x M$ and $\mathbb{R}^n \times \{0\}$, $\left|p_{T_x M} - p_{\mathbb{R}^n \times \{0\}}\right|^2$ is just $2\sum_{i,j}((e^{ij})^2 + (\varepsilon^{ij})^2 - 2e^{ij}\varepsilon^{ij}) = 2(n - \sum_{j=1}^n e^{jj}) = 2\sum_{j=n+1}^{n+\ell} e^{jj} = 2\sum_{j=1}^k |\nabla^M x^{n+j}|^2$, where we used the facts that $(e^{ij})^2 = (e^{ij})$ and $\operatorname{trace}(e^{ij}) = n$. Thus

$$\left|p_{T_x M} - p_{\mathbb{R}^n \times \{0\}}\right|^2 = 2\sum_{j=1}^k |\nabla^M x^{n+j}|^2, \tag{4.2}$$

[2] $\left|p_{T_x M} - p\right|^2$ denotes the inner product norm trace $(p_{T_x} - p)^2$; this differs from $\|p_{T_x M} - p\|^2$ by at most a constant factor depending on $n + \ell$—see Remark 4.4.

so in this case

4.3 $$E(\xi,\rho,T) = 2\sigma^{-n}\int_{B_\sigma(\xi)} \textstyle\sum_{j=1}^k \left|\nabla^M x^{n+j}\right|^2 d\mu_V$$

(∇^M = gradient operator on M as defined in §2 of Ch. 3).

4.4 Remark (Operator norm v. inner product norm): Notice that in 4.1 we use the inner product norm, but we could equivalently use the operator norm: If $L : \mathbb{R}^P \to R^Q$ is linear with matrix $\ell = (\ell_i^j)$ (so that $L(x) = \sum_{j=1}^Q \sum_{i=1}^P \ell_i^j x^i e_j$) then the operator norm is $\|L\| = \sup_{|x|=1} |L(x)|$, whereas the inner product norm is $|L| = \sqrt{\sum_{i,j} (\ell_i^j)^2}$. Observe $|L(x)|^2 = x^T \ell^T \ell x$ and $\ell^T \ell$ is a symmetric positive semi-definite $P \times P$ matrix with non-negative eigenvalues $0 \le \lambda_1 \le \lambda_2 \le \cdots \le \lambda_P$ and $|L|^2 = \operatorname{trace} \ell^T \ell = \sum_{j=1}^P \lambda_j$, while $\|L\|^2 = \lambda_P = \max\{\lambda_1, \ldots, \lambda_P\}$, so

$$P^{-1}|L|^2 \le \|L\|^2 \le |L|^2.$$

In particular $(n+\ell)^{-1} \int_{B_r(\xi)} |p_{T_xM} - p_{\mathbb{R}^n\times\{0\}}|^2 \le \int_{B_\sigma(\xi)} \|p_{T_xM} - p_{\mathbb{R}^n\times\{0\}}\|^2 \le \int_{B_\sigma(\xi)} |p_{T_xM} - p_{\mathbb{R}^n\times\{0\}}|^2$—i.e. $\int_{B_\sigma(\xi)} \|p_{T_xM} - p_{\mathbb{R}^n\times\{0\}}\|^2$ and $\int_{B_\sigma(\xi)} |p_{T_xM} - p_{\mathbb{R}^n\times\{0\}}|^2$ differ by at most fixed factor depending on $n+\ell$.

We are now ready to discuss the following Tilt-excess Decay Theorem, which is the main result concerning tilt-excess needed for the regularity theorem of the next section. In this theorem use the notation

$$E_*(\xi,\sigma,T) = \max\Big\{E(\xi,\sigma,T), \delta^{-1}\Big(\sigma^{p-n}\int_{B_\sigma(\xi)} \left|\underline{H}\right|^p d\mu\Big)^{2/p}\Big\},$$

where δ is as in 2.3.

4.5 Theorem (Tilt-excess Decay Theorem.) *There are constants $\eta, \delta_0 \in (0, \frac14]$, depending only on n, ℓ, p, such that if hypotheses 2.3 hold for some $\delta \in (0, \delta_0]$, if $\sigma \in (0, \delta\rho/2]$, $\xi \in \operatorname{spt}\mu_V \cap B_{\delta\rho}(0)$, and if T is any n-dimensional subspace of $\mathbb{R}^{n+\ell}$, then*

$$E_*(\xi,\eta\sigma,S) \le \eta^{2(1-n/p)} E_*(\xi,\sigma,T)$$

for some n-dimensional subspace $S \subset \mathbb{R}^{n+\ell}$.

4.6 Remark: Notice that any such S automatically satisfies

(‡) $$\left|p_S - p_T\right|^2 \le C\eta^{-n} E_*(\xi,\sigma,T)$$

Indeed we trivially have

$$(\eta\sigma)^{-n}\int_{B_{\eta\sigma}(\xi)} \left|p_{T_xM} - p_T\right|^2 d\mu \le \eta^{-n} E(\xi,\sigma,T),$$

while by 4.5 we have

$$(\eta\sigma)^{-n}\int_{B_{\eta\sigma}(\xi)} \left|p_{T_xM} - p_S\right|^2 d\mu \le E_*(\xi,\sigma,T),$$

and hence, since $\left|p_S - p_T\right|^2 \le 2\left|p_{T_xM} - p_T\right|^2 + 2\left|p_{T_xM} - p_S\right|^2$, (‡) follows by adding these inequalities and using the fact that $\mu(B_{\eta\sigma}(\xi)) \ge \frac12(\omega_n\eta\sigma)^n$ (by 2.4).

Proof of 4.5: Throughout the proof, $C = C(n,\ell,p)$. We can suppose $\xi = 0$, $T = \mathbb{R}^n \times \{0\}$. Observe that by the Affine Approximation Lemma 2.5 we have a subspace $\widetilde{T}$ (in fact $\widetilde{T} = T(0, 2\sigma)$ in the notation of 2.5) with $E(0,\sigma,\widetilde{T}) \le C\delta^{1/(n+1)}$, so we can certainly assume

(1) $$E(0,\sigma,T) \le C\delta^{1/(n+1)}$$

because otherwise we just prove the lemma with $\widetilde{T}$ in place of T and this then trivially implies the lemma for the original T. Notice that since $E(0,\sigma,\widetilde{T}) \le C\delta^{1/(n+1)}$ and $\sup_{x\in\operatorname{spt}\mu\cap B_\sigma(0)} \operatorname{dist}(x,\widetilde{T}) \le C\delta^{1/(2n+2)}$ by Lemma 2.5 we see also from (1) that $|p_T - p_{\widetilde{T}}| \le C\delta^{1/(2n+2)}$, and hence $\sup_{x\in\operatorname{spt}\mu\cap B_\sigma(0)} \operatorname{dist}(x,T) \le C\delta^{1/(2n+2)}$, which can be written

(2) $$\sup_{B_\sigma(0)\cap\operatorname{spt}\mu} \sum_{j=1}^{\ell} |x^{n+j}| \le C\delta^{1/(2n+2)}\sigma.$$

By the Lipschitz Approximation Lemma 2.6 with $L = 1$, there is a Lipschitz function $f : B_\sigma^n(0) \to \mathbb{R}^k$ with

(3) $$\operatorname{Lip} f \le 1,\ \sup|f| \le C\delta^{1/(2n+2)}\sigma$$

(4) $$\mu(\operatorname{spt}\mu \cap B_\sigma(0) \setminus \operatorname{graph} f) + \mathcal{H}^n(\operatorname{graph} f \cap B_\sigma(0) \setminus \operatorname{spt}\mu) \le CE_0\sigma^n,$$

where $E_0 = E_*(0,\sigma,\mathbb{R}^n\times\{0\})$; i.e.

$$E_* = \max\Big\{\sigma^{-n}\int_{B_\sigma(0)} \left|p_{T_xM} - p_{\mathbb{R}^n\times\{0\}}\right|^2 d\mu,\ \delta^{-1}\Big(\sigma^{p-n}\int_{B_\sigma(0)} \left|\underline{H}\right|^p d\mu\Big)^{2/p}\Big\}$$

Let us agree that $C\delta^{1/(2n+2)} \le 1/4$, in which case (2) implies

(5) $$(B_{\sigma/2}^n(0) \times \mathbb{R}^\ell) \cap \operatorname{spt}\mu \subset B_{\sigma/2}^n(0) \times B_{\sigma/4}^\ell(0).$$

Our aim now is to prove that each component of the Lipschitz function f is well-approximated by a harmonic function. Preparatory to this, note that the defining identity for $\underline{H}$ (see 3.14 of Ch. 4), with $X = \zeta e_{n+j}$, implies

$$\int_M \nabla_{n+j}^M \zeta\, d\mu = -\int e_{n+j}\cdot \underline{H}\zeta\, d\mu, \quad \zeta \in C_0^1(\breve{B}_\sigma(0)),$$

$j = 1,\ldots,\ell$, where $\nabla^M_{n+j} = e_{n+j}\cdot\nabla^M = p_{T_xM}(e_{n+j})\cdot\nabla^M = (\nabla^M x^{n+j})\cdot\nabla^M$ (∇^M = gradient operator for M as in §2 of Ch. 3). Thus we can write

$$\text{(6)}\qquad \int_M (\nabla^M x^{n+j})\cdot\nabla^M\zeta\,d\mu = -\int_M e_{n+j}\cdot\underline{H}\zeta\,d\mu.$$

Since $x^{n+j} \equiv \widetilde{f}^j(x)$ on $M_1 = M\cap\operatorname{graph} f$ (where $\widetilde{f}^j$ is defined on $\mathbb{R}^{n+\ell}$ by $\widetilde{f}^j(x^1,\ldots,x^{n+\ell}) = f^j(x^1,\ldots,x^n)$ for $x = (x^1,\ldots,x^n)\in\mathbb{R}^{n+\ell}$), we have by the definition of ∇^M (see §2 of Ch. 3) that

$$\text{(7)}\qquad \nabla^M x^{n+j} = \nabla^M\widetilde{f}^j(x)\quad \mu\text{-a.e. } x\in M_1 = M\cap\operatorname{graph} f.$$

Hence by (6) can be written

$$\int_{M_1}\nabla^M\widetilde{f}^j\cdot\nabla^M\zeta\,d\mu = -\int_{M\setminus M_1}(\nabla^M x^{n+j})\cdot\nabla^M\zeta\,d\mu - \int_M e_{n+j}\cdot\underline{H}\zeta\,d\mu,$$

and hence by (4), together with the fact that (by 2.3)

$$\int_{B_\sigma(\xi)}|\underline{H}|\,d\mu \le \Big(\int_{B_\sigma(\xi)}|\underline{H}|^p\,d\mu\Big)^{1/p}\big(\mu(B_\sigma(\xi))\big)^{1-1/p} \le C\delta^{1/2}E_0^{\frac12}\sigma^{n-1},$$

we obtain

$$\text{(8)}\qquad \Big|\sigma^{-n}\int_{M_1}(\nabla^M\widetilde{f}^j)\cdot\nabla^M\zeta\,d\mu\Big| \le C\big(\sigma^{-1}\sup|\zeta|\delta^{1/2}E_0^{\frac12} + \sup|\nabla\zeta|E_0\big)$$
$$\le C\sup|\nabla\zeta|(\delta^{1/2}E_0^{\frac12} + E_0),$$

for any smooth ζ with $\operatorname{spt}\zeta\subset\breve{B}_\sigma(0)$.

Furthermore by (7), 4.3, we evidently have

$$\text{(9)}\qquad \sigma^{-n}\int_{M_1\cap B_\sigma(0)}\big|\nabla^M\widetilde{f}^j\big|^2\,d\mu \le E_0.$$

Now suppose that ζ_0 is an arbitrary $C^1_c(B^n_{\sigma/2}(0))$ function, and let $\widetilde{\zeta}_0(x^1,\ldots,x^{n+\ell}) = \zeta_0(x^1,\ldots,x^n)$, so $\operatorname{spt}\widetilde{\zeta}_0 = \operatorname{spt}\zeta_0\times\mathbb{R}^\ell \subset \breve{B}^n_{\sigma/2}(0)\times\mathbb{R}^\ell$, and by (5) we can select a function $h\in C^1_c(B_\sigma(0))$ with $h\equiv 1$ in a neighborhood of $\operatorname{spt}\mu\cap\operatorname{spt}\widetilde{\zeta}_0$, and hence it is legitimate to use $\widetilde{\zeta}_0 h$ in place of ζ in the above discussion. But of course this is the same as using $\widetilde{\zeta}_0$ in place of ζ, since again $h\equiv 1$ in a neighborhood of $\operatorname{spt}\mu\cap\operatorname{spt}\widetilde{\zeta}_0$. So we can use all the above identities with $\widetilde{\zeta}_0$ in place of ζ. In particular (8) holds with $\widetilde{\zeta}_0$ in place of ζ. Also, since $p_{\mathbb{R}^n\times\{0\}}(\nabla\widetilde{\zeta}_0) = \nabla\widetilde{\zeta}_0$ and $p_{\mathbb{R}^n\times\{0\}}(\nabla\widetilde{f}^j) = \nabla\widetilde{f}^j$, we have

$$\text{(10)}\qquad \nabla^M\widetilde{f}^j\cdot\nabla^M\widetilde{\zeta}_0 = p_{T_xM}(\nabla\widetilde{f}^j)\cdot\nabla\widetilde{\zeta}_0$$
$$= \nabla\widetilde{f}^j\cdot\nabla\widetilde{\zeta}_0 - p_{(T_xM)^\perp}(\nabla\widetilde{f}^j)\cdot\nabla\widetilde{\zeta}_0$$
$$= \nabla\widetilde{f}^j\cdot\nabla\widetilde{\zeta}_0 - p_{(T_xM)^\perp}(\nabla\widetilde{f}^j)\cdot\nabla\widetilde{\zeta}_0$$
$$= \nabla\widetilde{f}^j\cdot\nabla\widetilde{\zeta}_0 - (p_{\mathbb{R}^n\times\{0\}}\circ p_{(T_xM)^\perp}\circ p_{\mathbb{R}^n\times\{0\}})(\nabla\widetilde{f}^j)\cdot\nabla\widetilde{\zeta}_0.$$

Now

$$\|p_{\mathbb{R}^n\times\{0\}}\circ p_{(T_xM)^\perp}\circ p_{\mathbb{R}^n\times\{0\}}\| = \|(p_{\mathbb{R}^n\times\{0\}} - p_{T_xM})\circ p_{(T_xM)^\perp}\circ(p_{\mathbb{R}^n\times\{0\}} - p_{T_xM})\|$$
$$\le \|p_{\mathbb{R}^n\times\{0\}} - p_{T_xM}\|^2 \le |p_{\mathbb{R}^n\times\{0\}} - p_{T_xM}|^2$$

by 4.4, so (10) implies

$$\text{(11)}\qquad \big|\nabla^M\widetilde{f}^j\cdot\nabla^M\widetilde{\zeta}_0\,d\mu - \nabla\widetilde{f}^j\cdot\nabla\widetilde{\zeta}_0\big| \le |p_{\mathbb{R}^n\times\{0\}} - p_{T_xM}|^2\sup|\nabla\widetilde{\zeta}_0|.$$

Thus (8) and (11) imply

$$\text{(12)}\qquad \Big|\sigma^{-n}\int_{M_1}\nabla\widetilde{f}^j\cdot\nabla\widetilde{\zeta}_0\,d\mu\Big| \le C(\delta^{1/2}E_0^{\frac12} + E_0)\sup|\nabla\zeta_0|.$$

Also since (10), (11) are valid with $\zeta_0 = f^j$, we conclude from (9) that

$$\text{(13)}\qquad \sigma^{-n}\int_{M_1\cap B_\sigma}\big|\nabla\widetilde{f}^j\big|^2\,d\mu \le CE_0.$$

From (12), (13) and the area formula 3.5 of Ch. 2 we then have (using also (3), (4))

$$\Big|\sigma^{-n}\int_{B^n_\sigma(0)}\nabla f^j\cdot\nabla\zeta_0\,\theta\circ F\,J_F\,d\mathcal{L}^n\Big| \le C\delta^{1/2}E_0^{1/2}\sup|\nabla\zeta_0|,$$

and

$$\text{(14)}\qquad \sigma^{-n}\int_{B^n_\sigma(0)}|\nabla f^j|^2\,\theta\circ F\,J_F\,d\mathcal{L}^n \le CE_0,$$

where $F:\mathbb{R}^n\to\mathbb{R}^{n+\ell}$ is the graph map defined by $x\in\mathbb{R}^n\mapsto F(x) = (x, f(x))\in\operatorname{graph} f\subset\mathbb{R}^{n+\ell}$, $x\in B^n_\sigma(0)$, and where J_F is the Jacobian of F defined as in §3 of Ch. 2 by

$$J_F(x) = \sqrt{\det\big(D_iF(x)\cdot D_jF(x)\big)} = \sqrt{\det\big(\delta_{ij} + D_if(x)\cdot D_jf(x)\big)}.$$

Then $1\le J_F\le 1 + C|\nabla f|^2$ on $B^n_\sigma(0)$ and $1\le\theta\le 1 + C\delta$ (by 2.5), so we conclude

$$\text{(15)}\qquad \Big|\sigma^{-n}\int_{B^n_\sigma(0)}\nabla f^j\cdot\nabla\zeta_0\,d\mathcal{L}^n\Big| \le C\Big(\delta^{\frac12}E_0^{\frac12} + \delta\sigma^{-n}\int_{B^n_{\sigma/2}(0)}|\nabla f^j|\,d\mathcal{L}^n\Big)\sup|\nabla\zeta_0|$$
$$\le C\delta^{1/2}E_0^{1/2}\sup|\nabla\zeta_0|$$

by (14), because by (14) (and the fact that $\theta\ge 1$, $J_F\ge 1$) we have

$$\text{(16)}\qquad \sigma^{-n}\int_{B^n_\sigma(0)}\big|\nabla f^j\big|^2\,d\mathcal{L}^n \le CE_0.$$

Now (15), (16) and the Harmonic Approximation 3.1 (with $(CE_0)^{-1/2}f^j$ in place of f) we know that for any given $\varepsilon\in(0,1)$ there is $\delta_0 = \delta_0(n)$ such that, if the hypotheses of 3.1 hold with $\delta \le \delta_0$, there are harmonic functions $u^1, \dots, u^\ell$ on $B_{\sigma/2}(0)$ such that

$$\text{(17)} \qquad \sigma^{-n}\int_{B^n_{\sigma/2}(0)} |Du|^2\, d\mathcal{L}^n \le CE_0, \;\; \sigma^{-n-2}\int_{B^n_{\sigma/2}(0)} |f-u|^2\, d\mathcal{L}^n \le \varepsilon E_0,$$

By (3) we have $|u(x)| \le |u(x) - f(x)| + C\delta^{1/(2n+2)}$ so $\int_{B_{\sigma/2}(0)} |u|^2 \le 2\int_{B_{\sigma/2}(0)} |u - f|^2 + C\delta^{1/(n+1)}E_0$, and hence by 3.4 and (17)

$$\text{(18)} \qquad \begin{cases} \sigma^{-1}|u(0)| \le C\left(\varepsilon^{1/2}E_0^{1/2} + \delta^{1/(2n+2)}\right) \le C\delta^{1/(2n+2)} \\ |Du(0)| \le CE_0^{1/2}. \end{cases}$$

Now, defining $\lambda(x) = (\lambda^1(x), \dots, \lambda^\ell(x))$ with $\lambda^j(x) = u^j(0) + x\cdot\nabla u^j(0)$ for $j = 1, \dots, \ell$, and again using 3.4 with $\eta \in (0, \frac14)$, we have also

$$\begin{aligned} \text{(19)} \quad (\eta\sigma)^{-n-2}\int_{B^n_{\eta\sigma}(0)} |f-\lambda|^2\, d\mathcal{L}^n &\le 2(\eta\sigma)^{-n-2}\int_{B_{\eta\sigma}(0)} \left(|f-u|^2 + |u-\lambda|^2\right) d\mathcal{L}^n \\ &\le 2\eta^{-n-2}\varepsilon E_0 + 2\omega_n\eta^{-2}\sigma^{-2}\sup\nolimits_{B_{\eta\sigma}(0)}|u-\lambda|^2 \\ &\le 2\eta^{-n-2}\varepsilon E_0 + C\eta^2\sigma^{-n}\int_{B^n_{\sigma}(0)} |Du|^2\, d\mathcal{L}^n \\ &\le 2\eta^{-n-2}\varepsilon E_0 + C\eta^2 E_0, \end{aligned}$$

where at the last step we used (17). Now let S be the n-dimensional subspace $\operatorname{graph}(\lambda - \lambda(0))$, let $\tau = (0, \lambda(0))$, and observe that $\operatorname{dist}(x, \tau + S) \le |f(x') - \lambda(x')|$ for any $x = (x', f(x')) \in B_{\eta\sigma}(\tau) \cap \operatorname{graph} f$, so (19) implies

$$(\eta\sigma)^{-n-2}\int_{B_{\eta\sigma}(\tau)\cap\operatorname{graph} f} \operatorname{dist}(x-\tau, S)^2\, d\mathcal{H}^n \le C\eta^{-n-2}\varepsilon E_0 + C\eta^2 E_0.$$

Then by (3), (4), (2), and (18), keeping in mind $\theta(\xi) \le 1 + C\delta \le 2$ in $B_\sigma(0)$,

$$(\eta\sigma)^{-n-2}\int_{B_{\eta\sigma}(\tau)} \operatorname{dist}(x-\tau, S)^2\, d\mu \le C\eta^{-n-2}\varepsilon E_0 + C\delta^{1/(n+1)}E_0 + C\eta^2 E_0,$$

and then by Remark 2.2 we have

$$\text{(20)} \qquad E(\tau, \eta\sigma/2, S) \le C\eta^{-n-2}\varepsilon E_0 + C\delta^{1/(n+1)}E_0 + C(\eta^2 + \delta)E_0.$$

Now (18) implies $|\tau| \le C\delta^{1/(2n+2)}\sigma$, hence

$$\text{(21)} \qquad C\delta^{1/(2n+2)} < \eta/4 \Rightarrow B_{\eta\sigma/4}(0) \subset B_{\eta\sigma/2}(\tau)$$

(for δ small enough depending on n, ℓ, p and η), and then (20) gives

$$\text{(22)} \qquad E(0, \eta\sigma/4, S) \le C\eta^{-n-2}(\varepsilon + \delta^{1/(2n+2)})E_0 + C(\eta^2 + \delta)E_0.$$

The proof is now completed as follows:

With C as in (22), first select $\eta = \eta(n, \ell, p)$ so that $C\eta^2 \le \frac14(\eta/4)^{2(1-n/p)}$, and then choose $\varepsilon = \varepsilon(n, \ell, p)$ so that $C\eta^{-n-2}\varepsilon \le \frac14(\eta/4)^{2(1-n/p)}$, and finally choose $\delta \le \delta_0(n, \ell, p)$ with δ_0 small enough so that $B_{\eta\sigma/4}(0) \subset B_{\eta\sigma/2}(\tau)$ as in (21) and so that the above harmonic approximation is valid with the choice of ε made above, and also so that $C\eta^{-n-2}\delta^{1/(2n+2)} \le \frac14(\eta/4)^{2(1-n/p)}$. Then (22) implies

$$\text{(23)} \qquad E(0, \tilde\eta\sigma, S) \le \tilde\eta^{2(1-n/p)}E_0,$$

where $\tilde\eta = \eta/4$. Since

$$\left((\tilde\eta\sigma)^{p-n}\int_{B_{\tilde\eta\sigma}(0)} |\underline{H}|^p\, d\mu\right)^{1/p} \le \tilde\eta^{1-n/p}\left(\sigma^{p-n}\int_{B_\sigma(0)} |\underline{H}|^p\, d\mu\right)^{1/p}$$

by virtue of the inclusion $B_{\tilde\eta\sigma}(0) \subset B_\sigma(0)$, we thus conclude that

$$E_*(0, \tilde\eta\sigma, S) \le \tilde\eta^{2(1-n/p)}E_*(0, \sigma, T).$$

This completes the proof of 4.5 (with $\tilde\eta$ in place of η). □

5 Main Regularity Theorem

We recall the hypotheses of §2 on V (which is a rectifiable varifold $V = \underline{v}(M, \theta)$ with generalized mean curvature $\underline{H}$ in the open set $U \subset \mathbb{R}^{n+\ell}$):

$$\textbf{5.1} \qquad \begin{cases} 1 \le \theta \;\mu\text{-a.e.},\; 0 \in \operatorname{spt} V,\; B_\rho(0) \subset U \\ \omega_n^{-1}\rho^{-n}\mu(B_\rho(0)) \le 1 + \delta,\; \left(\rho^{p-n}\int_{B_\rho(0)} |\underline{H}|^p\, d\mu\right)^{1/p} \le \delta. \end{cases}$$

Then we have the following:

5.2 Theorem (Allard Regularity Theorem.) *If $p > n$ is arbitrary, then there are $\delta_0 = \delta_0(n, \ell, p)$, $\gamma = \gamma(n, \ell, p) \in (0, 1)$ such that the hypotheses* 5.1 *with $\delta \le \delta_0$ imply the existence of a linear isometry q of $\mathbb{R}^{n+\ell}$ and a $u = (u^1, \dots, u^\ell) \in C^{1,1-n/p}(B^n_{\gamma\rho}(0); \mathbb{R}^\ell)$ with $Du(0) = 0$, $\operatorname{spt} V \cap B_{\gamma\rho}(0) = q(\operatorname{graph} u) \cap B_{\gamma\rho}(0)$, and*

$$\rho^{-1}\sup|u| + \sup|Du| + \rho^{1-n/p}\sup_{x,y\in B^n_{\gamma\rho}(0),\, x\ne y} |x-y|^{-(1-n/p)}|Du(x) - Du(y)| \le C\delta^{1/(2n+2)},$$

where $C = C(n, \ell, p)$.

5.3 Remark: We shall prove in the next section a slight improvement on the above theorem, in that *for every* $\gamma \in (0,1)$ there is $\delta = \delta(\gamma, n, \ell, p) \in (0,1)$ such that the hypotheses 5.1 imply the conclusion of the above theorem.

Proof: The proof is based on the Tilt-excess Decay 4.5 of the previous section. Throughout the proof $C = C(n, \ell, p) > 0$.

Take $\xi \in B_{\delta\rho/2}(0) \cap \operatorname{spt} V$ and $\sigma \in (0, \delta\rho/2]$ and let S_0 be an arbitrary n-dimensional subspace of $\mathbb{R}^{n+\ell}$. By the Tilt-excess Decay Theorem 4.5 we then know that there are $\delta = \delta(n, \ell, p)$, $\eta = \eta(n, \ell, p)$ so that 5.1 implies

$$E_*(\xi, \eta\sigma, S_1) \le \eta^{2(1-n/p)} E_*(\xi, \sigma, S_0)$$

for suitable S_1. Notice that this can be repeated; by induction we prove that if $\xi \in \operatorname{spt} V \cap B_{\delta\rho/2}(0)$, then, with $\sigma_0 = \delta\rho/2$, there is a sequence $S_1, S_2, \ldots$ of n-dimensional subspaces such that

$$(1) \quad E_*(\xi, \eta^j\sigma_0, S_j) \le \eta^{2(1-n/p)} E_*(\xi, \eta^{j-1}\rho/2, S_{j-1}) \le \eta^{2(1-n/p)j} E_*(\xi, \sigma_0, S_0)$$

for each $j \ge 1$.

Let $T_0 = T(0, 2\sigma_0)$; then 2.5 tells us that $E(0, \sigma_0, T_0) \le C\delta^{1/(n+1)}$ and hence, with the same C, $E(\xi, \sigma_0/2, T_0) \le 2^n C\delta^{1/(n+1)}$ for each $\xi \in \operatorname{spt}\mu \cap B_{\sigma_0/2}(0)$, so then the above, always taking $S_0 = T_0$ (for each $\xi \in B_{\sigma_0/2}(0) \cap \operatorname{spt}\mu$) implies

$$(2) \qquad E_*(\xi, \eta^j\sigma_0, S_j) \le \eta^{2(1-n/p)} E_*(\xi, \eta^{j-1}\sigma_0/2, S_{j-1}) \le \eta^{2(1-n/p)j} E_0,$$

where, here and subsequently, $E_0 = E_*(0, \sigma_0, T_0)$. Notice in particular that this gives (Cf. 4.6)

$$(3) \qquad \left|p_{S_j} - p_{S_{j-1}}\right|^2 \le C E_*(\xi, \eta^{j-1}\sigma_0, S_{j-1}) \le C\eta^{2(1-n/p)j} E_*(\xi, \sigma_0, S_0).$$

for each $j \ge 1$.

By summation from $j+1$ to ℓ, (3) gives

$$(4) \qquad \left|p_{S_\ell} - p_{S_j}\right|^2 \le C\eta^{2(1-n/p)j} E_0$$

for $\ell \ge j \ge 0$. (4) evidently implies that there is $S(\xi) (= \lim_{\ell\to\infty} S_\ell)$ such that

$$(5) \qquad \left|p_{S(\xi)} - p_{S_j}\right|^2 \le C\eta^{2(1-n/p)j} E_0, \quad j = 0, 1, 2, \ldots.$$

In particular (setting $j = 0$)

$$(6) \qquad \left|p_{S(\xi)} - p_{T_0}\right|^2 \le C E_0.$$

Now if $\sigma \in (0, \sigma_0/2]$ is arbitrary we can choose $j \ge 0$ such that $\eta^j\sigma_0/2 < \sigma \le \eta^{j-1}\sigma_0/2$. Then (1) and (5) evidently imply

$$(7) \qquad E_*(\xi, \sigma, S(\xi)) \le C(\sigma/\sigma_0)^{2(1-n/p)} E_0, \quad C = C(n, \ell, p),$$

for each $\xi \in B_{\sigma_0/2}(0) \cap \operatorname{spt} V$ and each $0 < \sigma \le \sigma_0/2$. Notice also that (6), (7) imply

$$(8) \qquad E_*(\xi, \sigma, T_0) \le CE_0 \le C\delta^{1/(2n+2)}, \quad 0 < \sigma \le \sigma_0/2.$$

Supposing without loss of generality that $T_0 = \mathbb{R}^n \times \{0\}$, we then see, by Corollary 2.7 and (8), if $L_0 \in (0, \frac{1}{4}]$ is given, and if $\delta \le \delta_0 L_0^{2n+2}$ for suitable $\delta_0 = \delta_0(n, \ell, p)$, then

$$(9) \qquad \operatorname{spt} V \cap B_{\sigma_0/4}(0) = \operatorname{graph} f \cap B_{\sigma_0/4}(0),$$

where f is a Lipschitz function $B^n_{\sigma_0/2}(0) \to \mathbb{R}^\ell$ with $\operatorname{Lip} f \le L_0$.

With such an f, let $G(f) = \operatorname{graph} f$ and $\xi = (\xi', f(\xi')) \in G(f)$, and note that, in view of (9), (7) implies

$$\lim_{\sigma\downarrow 0} \sigma^{-n} \int_{B_\sigma(\xi)\cap G(f)} |p_{T_xG(f)} - p_{S(\xi)}|^2 \, d\mathcal{H}^n = 0$$

for $\mathcal{H}^n$-a.e. $\xi \in G(f) \cap B_{\sigma_0/2}(0)$, and at all such points ξ it evidently follows that $S(\xi)$ is the approximate tangent space of $G(f)$; i.e. $S(\xi) = p_{T_\xi}G_f$, so (7) can be equivalently written

$$(10) \qquad \sigma^{-n} \int_{B_\sigma(\xi)\cap G(f)} |p_{T_xG(f)} - p_{T_\xi G(f)}|^2 \, d\mathcal{H}^n \le C(\sigma/\sigma_0)^{2(1-n/p)} E_0$$

for all $0 < \sigma \le \sigma_0/2$. Now the orthogonal projection $p_{T_\xi G(f)}$ of $\mathbb{R}^{n+\ell}$ onto the subspace $T_\xi G(f)$ is given by $p_{T_\xi G(f)}(v) = \sum_{j=1}^n (\tau_j \cdot v)\tau_j$, where τ_j is an orthonormal basis for $T_\xi G(f)$, and by the Gram-Schmidt orthogonalization process (starting with the basis $(e_j, D_j f(\xi'))$, $j = 1, \ldots, n$, for $T_\xi G(f)$, where $(\xi', f(\xi')) = \xi$) shows that $p_{T_\xi G(f)}$ has a matrix P_ξ of the form

$$P_\xi = \begin{pmatrix} I_{n\times n} & Df(\xi') \\ (Df(\xi'))^t & O_{\ell\times\ell} \end{pmatrix} + \mathcal{F}(Df(\xi')),$$

where $\mathcal{F}(p)$ is a real analytic function of $p = (p_{ij})_{i=1,\ldots,n, j=1,\ldots,\ell} \in \mathbb{R}^{n\ell}$ with $\mathcal{F}(0) = 0$, $D_p\mathcal{F}(0) = 0$ and hence $|\mathcal{F}(p_1) - \mathcal{F}(p_2)| \le C(n,\ell)(|p_1| + |p_2|)|p_1 - p_2|$ for $|p_1|, |p_2| \le 1$. Evidently then (provided we choose L_0 small enough, depending only on n, ℓ) we have

$$|Df(x') - Df(\xi')|^2 \le |p_{T_xG(f)} - p_{T_\xi G(f)}|^2 \le 3|Df(x') - Df(\xi')|^2$$

and so (10) implies

$$(11)\qquad \sigma^{-n}\int_{B^n_\sigma(\xi')}\big|Df(x)-Df(\xi)\big|^2\,d\mathcal{L}^n(x) \le C(\sigma/\sigma_0)^{2(1-n/p)}E_0,\quad 0<\sigma<\sigma_0/4.$$

For μ-a.e. $x_1,x_2\in \operatorname{spt} V\cap B_{\sigma_0/8}(0)$ we can use (11) with $\sigma=|x_1-x_2|$ and with $\xi = x_1,x_2$. Since $|Df(x_1)-Df(x_2)|^2\le 2|Df(x)-Df(x_1)|^2+2|Df(x)-Df(x_2)|^2$ for $x\in B^n_\sigma(x_1)\cap B^n_\sigma(x_2)\supset B^n_{\sigma/2}((x_1+x_2)/2)$ we then conclude

$$\big|Df(x_1)-Df(x_2)\big|\le C\left(|x_1-x_2|/\sigma_0\right)^{1-n/p}E_0^{1/2}$$

for $\mathcal{L}^n$-a.e. $x_1,x_2\in B^n_{\sigma_0/4}(0)$. Since f is Lipschitz it follows from this that $f\in C^{1,1-n/p}$ and this holds for *every* $x_1,x_2\in B^n_{\sigma_0/4}(0)$. Thus, choosing suitable $\delta=\delta(n,\ell,p)$ to satisfy the smallness restrictions imposed in the above argument, the theorem is established with $u=f$ and $\gamma=\delta/4$. □

6 Conical Approximation, Extension of Allard's Th.

First we want to derive an important technical lemma concerning conical approximation of V in the annular region $B_1(0)\setminus B_\lambda(0)$ in case $\lambda\in(0,\frac14)$ and $\omega_n^{-1}\mu_V(B_1(0))$ is close to $\omega_n^{-1}\lambda^{-n}\mu_V(B_\lambda(0))$.

6.1 Theorem. *Suppose $\Lambda>0$, $\lambda,\theta,\delta\in(0,\frac14)$, $\xi\in\partial B_{1-\theta}(0)$, $V=\underline{v}(M,\theta)$ has generalized mean curvature $\underline{H}\in L^1_{\mathrm{loc}}(\mu)$ in an open set $U\supset B_1(0)$, $\mu=\mu_V$, and*

$$\Big|\frac{\mu(B_1(0))}{\omega_n}-\frac{\mu(B_\lambda(0))}{\omega_n\lambda^n}\Big|+\sup_{r\in[\lambda,1]} r^{1-n}\int_{B_r(0)}|\underline{H}|\,d\mu\le\delta,$$

and $\mu(B_1(0))\le\Lambda$. Then

$$\mu(B_{(1-\delta^{1/4})\rho}(\xi))-C\delta^{1/4}|\log\lambda|\le\frac{\mu(B_{\tau\rho}(\tau\xi))}{\tau^n}\le\mu(B_{(1+\delta^{1/4})\rho}(\xi))+C\delta^{1/4}|\log\lambda|$$

for all $\rho\in(0,\theta]$ and all $\tau\in[2\lambda,1]$, where $C=C(n,\ell,\Lambda)$.

6.2 Remark: Note in particular that if $\underline{H}=0$ and if $\omega_n^{-1}\mu_V(B_1(0))=\Theta(\mu_V,0)$, then by the monotonicity identity 3.6 of Ch. 4 we would have $\omega_n^{-1}\mu_V(B_1(0)) = \omega_n^{-1}\lambda^{-n}\mu_V(B_\lambda(0))$ and hence the above theorem (with $\delta\downarrow 0$) guarantees that the density $\Theta(\mu_V,\tau\xi)$ is a constant function of $\tau\in(0,1)$ for each $\xi\in\partial B_1(0)$, so $V\llcorner\breve{B}_1(0)$ is a *cone*: $(\eta_{0,\tau\#}V)\llcorner\breve{B}_1(0)=V\llcorner\breve{B}_1(0)$ for each $\tau\in(0,1]$. This does indeed directly follow from the constancy of $\Theta^n(\mu_V,x)$ on the rays $R_\xi=\{x: x=\tau\xi:\tau\in(0,1)\}$, because, by Remark 4.10 of Ch. 4, we can write $V\llcorner\breve{B}_1(0)=\underline{v}(S\cap\breve{B}_1(0),\Theta)$, where $S=\operatorname{spt}\mu_V$ and $\Theta(x)=\Theta^n(\mu_V,x)$.

Proof of 6.1: First note that by 3.18 of Ch. 4 we have

$$(1)\qquad \int_{B_1(0)\setminus B_\lambda(0)} r^{-n-2}|D^\perp r|^2\,d\mu\le C\delta|\log\lambda|,$$

and then another application of 3.18 of Ch. 4 gives

$$(2)\qquad \Big|\frac{\mu(B_1(0))}{\omega_n}-\frac{\mu(B_\tau(0))}{\omega_n\tau^n}\Big|\le C\delta|\log\lambda|$$

for every $\tau\in[\lambda,1]$. Also, the hypothesis that $r^{1-n}\int_{B_r(0)}|\underline{H}|\le\delta$ for $r\in[\lambda,1]$, together with Lemma 4.11 of Ch. 4, implies

$$(3)\qquad \int_{B_1(0)\setminus B_\lambda(0)} r^{1-n}|\underline{H}|\,d\mu\le C\delta|\log\lambda|.$$

Let $h:\mathbb{R}^{n+\ell}\setminus\{0\}\to[0,1]$ be a homogeneous degree zero C^1 function on $\mathbb{R}^{n+\ell}\setminus\{0\}$ with $|Dh(x)|\le k/|x|$, where $k\ge1$ will be chosen below. Observe that $x\cdot p_{T_xM}(Dh(x)) = p_{T_xM}(x)\cdot Dh(x)=(x-p_{(T_xM)^\perp}(x))\cdot Dh(x)=-p_{(T_xM)^\perp}(x)\cdot Dh(x)$, so in fact

$$(4)\qquad |x\cdot p_{T_xM}(Dh(x))|\le k\rho^{-1}|x|^{-1}|p_{(T_xM)^\perp}(x)|.$$

We now let $0<s<t<1$, $h_1\in(0,s)$, $h_2\in(0,1-t)$ and use the first variation identity $\int \operatorname{div}_M X\,d\mu=-\int X\cdot\underline{H}\,d\mu$ with $X_{|x}=h(x)r^{-n}\gamma(r)x$, where $r=|x|$ and γ is defined by

$$\gamma(r)=\begin{cases}1 & \text{if } r\in[s,t]\\ 0 & \text{if } r<s-h_1 \text{ or } r>t+h_2\end{cases}$$

and $\gamma(r)$ is piecewise linear with slope h_1^{-1} for $r\in[s-h_1,s]$ and slope h_2^{-1} for $r\in[t,t+h_2]$. After letting $h_1,h_2\downarrow0$ (in any order) and observing that $\operatorname{div}_M(r^{-n}x)=nr^{-n}|p_{(T_xM)^\perp}(x)|^2$ we see that this gives

$$\frac{d}{ds}\int_{B_s(0)}h(x)r^{1-n}|\nabla^M r|^2\,d\mu-\frac{d}{dt}\int_{B_t(0)}h(x)r^{1-n}|\nabla^M r|^2\,d\mu + n\int_{B_t(0)\setminus B_s(0)}h(x)r^{-n-2}|p_{T_x^\perp}(x)|^2=-\int_{B_t(0)\setminus B_s(0)}h(x)r^{-n}x\cdot(\underline{H}-Dh(x))\,d\mu,$$

where $T_x^\perp=(T_xM)^\perp$. Since

$$\frac{d}{dt}\int_{B_t(0)}h(x)r^{-n}|\nabla^M r|^2\,d\mu=\int_{\partial B_t(0)}h(x)r^{-n}|\nabla^M r|\,\theta\,d\mathcal{H}^{n-1}$$

(by the co-area formula 2.9 of Ch.3) we see that, after using (3) and (4), this implies

$$(5)\qquad \Big|\int_{\partial B_s(0)}hr^{1-n}|\nabla^M r|-\int_{\partial B_t(0)}hr^{1-n}|\nabla^M r|+n\int_{B_t(0)\setminus B_s(0)}hr^{-n-2}|p_{T_x^\perp}(x)|^2\Big| \le C\delta|\log\lambda|+k\rho^{-1}\int_{B_t(0)\setminus B_s(0)}r^{-n-1}|p_{T_x^\perp}(x)|\,d\mu.$$

Now we let $s = \tau t$, $\tau \in [\lambda, \frac{1}{2}]$, thus giving

(6) $$\Big|\tau^{1-n}\int_{\partial B_{\tau t}(0)} h|\nabla^M r| - \int_{\partial B_t(0)} h|\nabla^M r|\Big| \le n\int_{B_1(0)\setminus B_\lambda(0)} hr^{-n-2}|p_{T_x^\perp}(x)|^2 + C\delta|\log\lambda| + k\rho^{-1}\int_{B_1(0)\setminus B_\lambda(0)} r^{-n-1}|p_{T_x^\perp}(x)|\,d\mu.$$

By using Cauchy-Schwarz and the bound $\mu(B_1(0)) \le \Lambda$ together with 3.18 of Ch. 4 in the last term on the right we infer

(7) $$\Big|\tau^{1-n}\int_{\partial B_{\tau t}(0)} h|\nabla^M r| - \int_{\partial B_t(0)} h|\nabla^M r|\Big| \le Ck\Lambda^{1/2}\Big(\int_{B_1(0)\setminus B_\lambda(0)} hr^{-n-2}|p_{T_x^\perp}(x)|^2\,d\mu\Big)^{1/2} + C\delta|\log\lambda|.$$

Now for any $\rho \in (0,]$ and $\xi \in \partial B_{1-\theta}(0)$ and for any $t \in (|\xi| - \rho, \xi + \rho)$ we select the homogeneous function h so that h is identically 1 on $\partial B_t(0) \cap B_{(1-\delta^{1/4})\rho}(\xi)$, identically zero on $\partial B_t(0) \setminus B_\rho(\xi)$, and $|Dh(x)| \le (4\delta^{1/4}\rho|x|)^{-1}$. Then (7) implies

(8)
$$\int_{\partial B_t(0)\cap B_{(1-\delta^{1/4})\rho}(\xi)} |\nabla^M r| \le \tau^{1-n}\int_{\partial B_{\tau t}(0)\cap B_{\tau\rho}(\tau\xi)} |\nabla^M r| \le C\delta^{-1/4}\rho^{-1}\Lambda^{1/2}|\log\lambda|^{1/2}\Big(\int_{B_1(0)\setminus B_\lambda(0)} r^{-n-2}|p_{T_x^\perp}(x)|^2\,d\mu\Big)^{1/2} + C\delta|\log\lambda|.$$

Then by integrating with respect to $t \in (|\xi|-\rho, |\xi|+\rho)$ (hence $\tau t \in (|\tau\xi|-\tau\rho, |\tau\xi|+\tau\rho)$), by the coarea formula we get

(9)
$$\int_{B_{(1-\delta^{1/4})\rho}(\xi)} |\nabla^M r|^2 \le \tau^{-n}\int_{B_{\tau\rho}(\tau\xi)} |\nabla^M r|^2 \le C\delta^{-1/4}\Lambda^{1/2}|\log\lambda|^{1/2}\Big(\int_{B_1(0)\setminus B_\lambda(0)} r^{-n-2}|p_{T_x^\perp}(x)|^2\,d\mu\Big)^{1/2} + C\delta\rho|\log\lambda|.$$

and since $|\nabla^M r|^2 = 1 - r^{-2}|p_{T_x^\perp}(x)|^2$ this implies

(10)
$$\mu(B_{(1-\delta^{1/4})\rho}(\xi)) \le \tau^{-n}\mu(B_{\tau\rho}(\tau\xi)) + C\int_{B_1(0)\setminus B_\lambda(0)} r^{-2}|p_{T_x^\perp}(x)|^2 \le C\delta^{-1/4}\Lambda^{1/2}|\log\lambda|^{1/2}\Big(\int_{B_1(0)\setminus B_\lambda(0)} r^{-n-2}|p_{T_x^\perp}(x)|^2\,d\mu\Big)^{1/2} + C\delta\rho|\log\lambda|.$$

Using (1), we then have

(11) $$\mu(B_{(1-\delta^{1/4})\rho}(\xi)) \le \tau^{-n}\mu(B_{\tau\rho}(\tau\xi)) + C\delta^{1/4}\Lambda^{1/2}|\log\lambda|,$$

with $C = C(n, \ell)$.

Similarly, choosing the homogeneous function h in (6) such that h is identically 1 on $\partial B_t(0) \cap B_\rho(\xi)$, identically zero on $\partial B_t(0) \setminus B_{(1+\delta^{1/4})\rho}(\xi)$, and $|Dh(x)| \le (4\delta^{1/4}\rho|x|)^{-1}$, we obtain

(12) $$\tau^{-n}\mu(B_{\tau\rho}(\tau\xi)) \le \mu(B_{(1+\delta^{1/4})\rho}(\xi)) + C\delta^{1/4}\Lambda^{1/2}|\log\lambda|. \qquad \square$$

7 Some Initial Applications of the Allard Theorem

The Allard Theorem of the previous two sections is fundamental in the study of the regularity and compactness properties of rectifiable varifolds (including also smooth submanifolds) with prescribed (generalized) mean curvature, in particular in the study of stationary varifolds. Here we discuss some initial applications. First we have the following corollary of the version of the Allard theorem discussed in the previous section:

7.1 Theorem. *If $V = \underline{v}(M, \theta)$, of dimension n, has generalized mean curvature $\underline{H}$ (as in 3.14 of Ch. 4) in an open set $U \subset \mathbb{R}^{n+\ell}$ and if $\underline{H}$ is locally in $L^p(\mu_V)$ for some $p > n$, if $\theta \ge 1$ μ_V-a.e. in U and if $\xi \in U$ with $\Theta^n(\mu_V, \xi) = 1$, then there is $\rho > 0$ and an orthogonal Q and of $\mathbb{R}^{n+\ell}$ such that, up to a set of $\mathcal{H}^n$ measure zero,*

$$Q \circ \tau(M) \cap \breve{B}_\rho(0) = \operatorname{graph} u, \quad \tau : x \mapsto x - \xi,$$

where $u : W \to \mathbb{R}^k$, W open in $\mathbb{R}^n$, is a $C^{1,1-n/p}(W, \mathbb{R}^k)$ function with $u(0) = 0, |Du(0)| = 0$.

In case θ is positive integer-valued μ_V-a.e. in U and $\underline{H} = h|\operatorname{spt} V$, where h is a $C^{q,\alpha}$ function in U for some $q \in \{0, 1, 2, \ldots\}$ and some $\alpha \in (0, 1)$, then, for sufficiently small $\rho > 0$, the above u is automatically $C^{q+2,\alpha}(W)$ and $\Theta^n(\mu_V, x) \equiv 1$ on $\breve{B}_\rho(\xi)$.

Finally, if θ is positive integer-valued μ_V-a.e. in U, if N is an $(n+\ell)$-dimensional C^{q+3} submanifold of $\mathbb{R}^{n+\ell}$ and if V is stationary in $U \cap N$ as in 2.7 of Ch. 4 (so that V has generalized mean curvature $\underline{H} = \overline{\underline{H}}_M$ as in 3.14 of Ch. 4), then again, for sufficiently small $\rho > 0$, the above u is automatically in $C^{q+2,\alpha}(W)$, this time for each $\alpha \in (0, 1)$, and $\Theta^n(\mu_V, x) \equiv 1$ in $\breve{B}_\rho(\xi) \cap \operatorname{spt}\mu_V$.

Proof: Since $\lim_{\rho\downarrow 0}\big(\rho^{p-n}\int_{B_\rho(\xi)} |H|^p\,d\mu_V\big)^{1/p} = 0$ and $\lim_{\rho\downarrow 0}(\omega_n\rho^n)^{-1}\mu_V(B_\rho(\xi)) = 1$, we can choose $\rho > 0$ such that the hypotheses of Theorem 5.2 hold, so, after applying the appropriate translation and orthogonal transformation, the required u exists with

(1) $$\operatorname{graph} u = \operatorname{spt} V \cap B_\sigma(0)$$

with $\sigma = \gamma\rho$, γ as in Theorem 5.2. Since θ is integer valued and < 2 a.e., we have $\theta = 1$ $\mathcal{H}^n$-a.e. on graph u; but graph u is a C^1 submanifold so then $\Theta^n(\mu_V, x) = 1$ at every point of graph u.

Let $\varepsilon_0 \in (0,1)$. Since $Du(0) = 0$, by choosing a smaller σ if necessary, we can assume that $|Du| \le \varepsilon_0$ on $B_\sigma(0)$ and so the analysis we made in §1 of the present chapter is applicable and tells us that u satisfies a system of equations of the form 1.3; i.e.

$$\text{(2)} \qquad \Delta u_i = \textstyle\sum_{j=1}^n D_j(A_{ij}(Du)) + h_i,\ i = 1,\dots,k,$$

with $A_{ij}(P)$ are C^∞ functions of the variable $P = (p_{\ell m})_{\ell=1,\dots,n,m=1,\dots,k}$ with $|A_{ij}(P)| \le C|P|^2$ and $|D_P A_{ij}(P)| \le C|P|$, where $C = C(n)$. Then by the Schauder theory for elliptic equations we see that $h_i \in C^{q,\alpha}(\breve{B}_\sigma(0))$ implies that $u \in C^{q+2,\alpha}(\breve{B}_\sigma(0))$ as claimed.

Finally, assume V is stationary in N. Then we can apply Theorem 5.2 for each $p > n$ so for each $\alpha \in (0,1)$ we have σ such that (1) holds with $u \in C^{1,\alpha}(\breve{B}_\sigma(0))$. Then $(e_i, D_i u(x)), i = 1,\dots,n$, is a $C^{0,\alpha}$ basis for $T_{(x,u(x))}G$, $G = \operatorname{graph} u$, $x \in \breve{B}_\sigma(0)$. By the Gram-Schmidt orthogonalization theorem we then have functions $F_j(Du)$, $j = 1,\dots,n$, such that $F_j(P)$ is a smooth function of $P = (p_{ij})_{i=1,\dots,n,j=1,\dots,k}$ and $F_1(Du(x)),\dots,F_n(Du(x))$ is an orthonormal basis for $T_{(x,u(x))}G$ for each $x \in \breve{B}_\sigma(0)$. Then, by 2.7 of Ch. 4, G has generalized mean curvature at $(x,u(x))$ equal to $\sum_{j=1}^n \overline{B}_{u(x)}(F_j(Du(x)), F_j(Du(x)))$. Thus, in this case (1) can be written

$$\text{(3)} \quad \Delta u_i = \textstyle\sum_{j=1}^n D_j(A_{ij}(Du)) + \sum_{j=1}^n e_{n+i}\cdot \overline{B}_{(x,u(x))}(F_j(Du(x)), F_j(Du(x)))$$

for $i = 1,\dots,k$, and again standard elliptic theory implies $u \in C^{q+2,\alpha}(\breve{B}_\sigma(0))$. $\square$

7.2 Definition: If $V = \underline{v}(M,\theta)$ is an n-dimensional rectifiable varifold, we say that a point $\xi \in \operatorname{spt} V$ is a *regular point* of V if there is a $\rho > 0$ such that $\breve{B}_\rho(\xi) \cap \operatorname{spt} V$ is an n-dimensional C^1 submanifold of $\mathbb{R}^{n+\ell}$. Then we let

$$\operatorname{reg} V = \{\xi \in \operatorname{spt} V : \xi \text{ is a regular point of } V\}$$
$$\operatorname{sing} V = \operatorname{spt} V \setminus \operatorname{reg} V.$$

Notice that then by definition reg V, sing V are respectively relatively open in spt V and relatively closed in U.

7.3 Corollary. *If $V = \underline{v}(M,\theta)$, of dimension n, has generalized mean curvature $\underline{H}$ in an open set $U \subset \mathbb{R}^{n+\ell}$, if $\underline{H}$ is locally in $L^p(\mu_V)$ for some $p > n$, and if θ is positive integer-valued μ_V-a.e. in* spt V, *then* reg V *is a relatively open dense set in* spt V*; i.e.* sing V *is nowhere dense in* spt V*, and* spt V *is the closure, taken in U, of* reg V.

7.4 Remark: It is an open question whether on not sing V has $\mathcal{H}^n$-measure zero under the general conditions of the above corollary, even if we assume $\underline{H} = 0$; such results (and more) are true in the special case when V is the varifold associated with a minimizing current, as discussed below in Ch. 7.

Proof of 7.3: Take any ball $B_\rho(\xi) \subset U$ and let

$$N = \min\{j : j \in \{1,2,\dots\} \text{ and } \Theta^n(\mu_V, x) = j \text{ for some } x \in \breve{B}_\rho(\xi)\}.$$

Then $\widetilde{V} = \underline{v}(M, N^{-1}\theta) \llcorner \breve{B}_\rho(\xi)$ has density $\Theta^n(\mu_{\widetilde{V}}, x) \ge 1$ everywhere in spt $\widetilde{V} \cap \breve{B}_\rho(\xi)$ and $\Theta^n(\mu_{\widetilde{V}}, x_0) = 1$ at some point of $x_0 \in \breve{B}_\rho(\xi)$. Such a point x_0 is in reg $\widetilde{V}$ $(=\operatorname{reg} V \cap \breve{B}_\rho(\xi))$ by Theorem 7.1, so we have shown reg $V \cap \breve{B}_\rho(\xi) \ne \emptyset$. $\square$

The Allard theorem will play a key role later (in Ch. 7) in establishing the regularity theory for solutions of the Plateau problem in arbitrary dimensions.

Chapter 6

Currents

1 Preliminaries: Vectors, Co-vectors, and Forms

$e_1, \dots, e_p$ denote the standard orthonormal basis for $\mathbb{R}^P$. We let $\Lambda^1(\mathbb{R}^P)$ denote the dual space of $\mathbb{R}^P$; thus $\Lambda^1(\mathbb{R}^P)$ is the space of linear functionals $\omega : \mathbb{R}^P \to \mathbb{R}$. $dx^1, \dots, dx^P \in \Lambda^1(\mathbb{R}^P)$ will denote the basis for $\Lambda^1(\mathbb{R}^P)$ dual to the standard basis $e_1, \dots, e_P$ of $\mathbb{R}^P$. Thus for $v = (v_1, \dots, v_P) \in \mathbb{R}^P$ we have

$$dx^j(v) = v_j, \;\; j = 1, \dots, P.$$

For $n \geq 2$, $\Lambda^n(\mathbb{R}^P)$ denotes the space of alternating n-linear functions on $\mathbb{R}^P \times \cdots \times \mathbb{R}^P$ (n factors). Thus $\omega \in \Lambda^n(\mathbb{R}^P)$ means $\omega(v_1, \dots, v_n)$ is linear in each v_j and $\omega(v_1, \dots v_i, \dots, v_j, \dots, v_n) = -\omega(v_1, \dots v_j, \dots, v_i, \dots, v_n)$ for each $i \neq j$. If $\omega_1, \dots, \omega_n \in \Lambda^1(\mathbb{R}^P)$ we define $\omega_1 \wedge \omega_2 \wedge \cdots \wedge \omega_n \in \Lambda^n(\mathbb{R}^P)$ by

1.1 $$\omega_1 \wedge \omega_2 \wedge \cdots \wedge \omega_n(v_1, \dots, v_n) = \sum_\sigma \operatorname{sgn}\sigma \; \omega_{\sigma(1)}(v_1)\omega_{\sigma(2)}(v_2)\cdots\omega_{\sigma(n)}(v_n) \;\; (= \det(\omega_i(v_j))),$$

where the sum is over all permutations σ of $\{1, \dots, n\}$ and where $\operatorname{sgn}\sigma$ is the sign of the permutation $\sigma : \{1, \dots, n\} \to \{1, \dots, n\}$. One easily checks that then any

$\omega \in \Lambda^n(\mathbb{R}^P)$ can be represented

$$\begin{aligned}\omega &= \textstyle\sum_{1\le i_1<\cdots<i_n\le P} \omega_{i_1\ldots i_n}\, dx^{i_1}\wedge\cdots\wedge dx^{i_n}\\ &= \textstyle\sum_{\alpha\in I_{n,P}} \omega_\alpha\, dx^\alpha\end{aligned}$$

where

1.2 $$I_{n,P} = \{\alpha = (i_1,\ldots,i_n) \in \mathbb{Z}_+^n : 1 \le i_1 < \cdots < i_n \le P\},$$

and where we use "multi-index" notation, with $\alpha = (i_1,\ldots,i_n) \in \mathbb{Z}_+^n$ ($\mathbb{Z}_+ =$ the set of non-negative integers), and where $\omega_{i_1,\ldots,i_n} = \omega(e_{i_1},\ldots,e_{i_n})$. Thus $\{dx^\alpha : \alpha \in I_{n,P}\}$ are a basis for $\Lambda^n(\mathbb{R}^P)$ and the dimension is $\binom{P}{n}$.

For $\omega = \sum_{\alpha\in I_{\ell,P}} \omega_\alpha dx^\alpha \in \Lambda^\ell(\mathbb{R}^P)$, $\eta = \sum_{\beta\in I_{m,P}} \eta_\beta dx^\beta \in \Lambda^m(\mathbb{R}^P)$ we can define

1.3 $$\omega\wedge\eta = \sum_{\alpha\in I_{\ell,P},\beta\in I_{m,P}} \omega_\alpha\eta_\beta\, dx^\alpha\wedge dx^\beta \in \Lambda^{\ell+m}(\mathbb{R}^P).$$

This is consistent with 1.1, and for $\omega,\omega_1,\omega_2 \in \Lambda^\ell(\mathbb{R}^P), \eta\in\Lambda^m(\mathbb{R}^P), \nu\in\Lambda^p(\mathbb{R}^P)$ we have

$$\begin{aligned}(c_1\omega_1 + c_2\omega_2)\wedge\eta &= c_1\omega_1\wedge\eta + c_2\omega_2\wedge\eta\\ (\omega\wedge\eta)\wedge\nu &= \omega\wedge(\eta\wedge\nu)\\ \omega\wedge\eta &= (-1)^{\ell m}\eta\wedge\omega.\end{aligned}$$

If V is a subspace of $\mathbb{R}^P$ of dim $= n$ with basis $v_1,\ldots,v_n$ then $\Lambda^n(V)$ denotes the subspace of $\Lambda^n(\mathbb{R}^P)$ with basis $\{v_{i_1}^*\wedge\cdots\wedge v_{i_n}^* : (i_1,\ldots,i_n)\in I_{n,\ell}\}$, where $v_j^*\in\Lambda^1(\mathbb{R}^P)$ is the element dual to v_j, so that, for $v\in\mathbb{R}^P$, $v^*\in\Lambda^1(\mathbb{R}^P)$ is defined by

1.4 $$v^*(w) = v\cdot w,\ \ w\in\mathbb{R}^P.$$

Analogous to the definition of $\Lambda^n(\mathbb{R}^P)$, we could similarly define $\Lambda^n(\Lambda^1(\mathbb{R}^P))$ for $n \ge 2$ as the space of alternating n-linear functions on $\Lambda^1(\mathbb{R}^P)$. In which case, after making the identification of $(dx^j)^* \simeq e_j$, we have the space $\Lambda_n(\mathbb{R}^P) \simeq \Lambda^n(\Lambda^1(\mathbb{R}^P))$ of n-vectors

$$w = \textstyle\sum_{\alpha\in I_{n,P}} w^\alpha e_\alpha,$$

where $w^\alpha\in\mathbb{R}$ and $e_\alpha = e_{j_1}\wedge\cdots\wedge e_{j_n}$ for $\alpha = (j_1,\ldots,j_n)\in I_{n,P}$, and

$$\begin{aligned}v_1\wedge\cdots\wedge v_n &= \textstyle\sum_{j_1,\ldots,j_n=1}^P v_{1j_1}v_{2j_2}\cdots v_{nj_n} e_{j_1}\wedge\cdots\wedge e_{j_n}\\ &= \textstyle\sum_{(\ell_1,\ldots,\ell_n)\in I_{n,P}} \det(v_{i\ell_j})\, e_{\ell_1}\wedge\cdots\wedge e_{\ell_n}\end{aligned}$$

for any $v_1,\ldots,v_n\in\mathbb{R}^P$.

If V is a subspace of $\mathbb{R}^P$ of dim $= n$ with basis $v_1,\ldots,v_n$ then $\Lambda_n(V)$ is the subspace of $\Lambda_n(\mathbb{R}^P)$ spanned by $\{v_{i_1}\wedge\cdots\wedge v_{i_n} : (i_1,\ldots,i_n)\in I_{n,\ell}\}$.

$\omega\in\Lambda^n(\mathbb{R}^P)$ (respectively $w\in\Lambda_n(\mathbb{R}^P)$) is called *simple* if it can be expressed $\omega_1\wedge\cdots\wedge\omega_n$ with $\omega_j\in\Lambda^1(\mathbb{R}^P)$ (respectively $w_1\wedge\cdots\wedge w_n$ with $w_j\in\mathbb{R}^P$).

We assume $\Lambda_n(\mathbb{R}^P)$, $\Lambda^n(\mathbb{R}^P)$ are equipped with the inner products naturally induced from $\mathbb{R}^P$ (making $\{e_\alpha\}_{\alpha\in I_{n,P}}$, $\{dx^\alpha\}_{\alpha\in I_{n,P}}$ orthonormal bases). Thus

1.5 $$\left(\textstyle\sum_{\alpha\in I_{n,P}}\omega_\alpha dx^\alpha\right)\cdot\left(\textstyle\sum_{\alpha\in I_{n,P}}\eta_\alpha dx^\alpha\right) = \textstyle\sum_{\alpha\in I_{n,P}}\omega_\alpha\,\eta_\alpha$$

and

1.6 $$\left(\textstyle\sum_{\alpha\in I_{n,P}}u^\alpha e_\alpha\right)\cdot\left(\textstyle\sum_{\alpha\in I_{n,P}}w^\alpha e_\alpha\right) = \textstyle\sum_{\alpha\in I_{n,P}}u^\alpha\, w^\alpha$$

The dual pairing between $\omega\in\Lambda^n(\mathbb{R}^P)$ and $w\in\Lambda_n(\mathbb{R}^P)$ will be denoted $\langle\omega,w\rangle$; thus

1.7 $$\left\langle \textstyle\sum_{\alpha\in I_{n,P}}\omega_\alpha dx^\alpha, \sum_{\alpha\in I_{n,P}} w^\alpha e_\alpha\right\rangle = \sum_{\alpha\in I_{n,P}}\omega_\alpha w^\alpha.$$

Given $\ell:\mathbb{R}^P\to\mathbb{R}^Q$ linear, the "pull-back" $\ell^\#:\Lambda^n(\mathbb{R}^Q)\to\Lambda^n(\mathbb{R}^P)$ is defined by

1.8 $$\ell^\#\omega(v_1,\ldots,v_n) = \omega(\ell(v_1),\ldots,\ell(v_n)),\ \ v_1,\ldots,v_n\in\mathbb{R}^Q,$$

and then the "push-forward" $\ell_\#:\Lambda^n(\mathbb{R}^P)\to\Lambda^n(\mathbb{R}^Q)$ is defined by duality according to the requirement

1.9 $$\langle\ell^\#\omega, w\rangle = \langle\omega,\ell_\# w\rangle,\quad \omega\in\Lambda^n(\mathbb{R}^Q),\ w\in\Lambda_n(\mathbb{R}^P),$$

where $\langle\,,\rangle$ is the dual pairing as in 1.7. More explicitly, $\ell^\#,\ell_\#$ are then characterized as the unique linear maps $\Lambda^n(\mathbb{R}^Q)\to\Lambda^n(\mathbb{R}^P)$ and $\Lambda_n(\mathbb{R}^P)\to\Lambda_n(\mathbb{R}^Q)$ respectively such that

1.10 $$\begin{cases}\ell^\#(\omega_1\wedge\cdots\wedge\omega_n) = (\omega_1\circ\ell)\wedge\cdots\wedge(\omega_n\circ\ell), & \omega_1,\ldots,\omega_n\in\Lambda^1(\mathbb{R}^Q)\\ \ell_\#(v_1\wedge\cdots\wedge v_n) = \ell(v_1)\wedge\cdots\wedge\ell(v_n), & v_1,\ldots,v_n\in\mathbb{R}^P.\end{cases}$$

For open $U\subset\mathbb{R}^P$, $\mathcal{E}^n(U) = C^\infty(U,\Lambda^n(\mathbb{R}^P))$ and the elements $\omega\in\mathcal{E}^n(U)$ are called smooth n-forms on U. Thus $\omega\in\mathcal{E}^n(U)$ means $\omega = \sum_{\alpha\in I_{n,P}}\omega_\alpha dx^\alpha$ where $\omega_\alpha\in C^\infty(U)$.

The value of $\omega(x) = \sum_{\alpha\in I_{n,P}}\omega_\alpha(x)dx^\alpha$ at a point $x\in U$ will also at times be denoted $\omega|_x$.

The exterior derivative $\mathcal{E}^n(U) \to \mathcal{E}^{n+1}(U)$ is defined as usual by

1.11 $$d\omega = \sum_{j=1}^{P} \sum_{\alpha \in I_{n,P}} \frac{\partial a_\alpha}{\partial x^j} dx^j \wedge dx^\alpha$$

if $\omega = \sum_{\alpha \in I_{n,P}} a_\alpha dx^\alpha$. By direct computation (using $\frac{\partial^2 a_\alpha}{\partial x^i \partial x^j} = \frac{\partial^2 a_\alpha}{\partial x^j \partial x^i}$ and $dx^i \wedge dx^j = -dx^j \wedge dx^i$) one checks that

1.12 $$d^2\omega = 0 \ \forall \omega \in \mathcal{E}^n(U).$$

Given $\omega = \sum_{\alpha \in I_{n,Q}} \omega_\alpha(y) dy^\alpha \in \mathcal{E}^n(V)$, $V \subset \mathbb{R}^Q$ open, and a smooth map $f : U \to V$, we define the "pulled back" form $f^\# \omega \in \mathcal{E}^n(U)$ by

1.13 $$f^\# \omega = \sum_{\alpha=(i_1,\ldots,i_n) \in I_{n,Q}} \omega_\alpha \circ f \, df^{i_1} \wedge \cdots \wedge df^{i_n},$$

where df^j is $\sum_{i=1}^{P} \frac{\partial f^j}{\partial x^i} dx^i$, $j = 1, \ldots, Q$. Equivalently this says

$$f^\# \omega|_x = (df_x)^\# (\omega|_{f(x)}),$$

where the right side is defined as in 1.8 with $\ell = df_x$.

Notice that the exterior derivative commutes with the pulling back:

1.14 $$df^\# = f^\# d.$$

We let $\mathcal{D}^n(U)$ denote the set of $\omega = \sum_{\alpha \in I_{n,P}} \omega_\alpha dx^\alpha \in \mathcal{E}^n(U)$ such that each ω_α has compact support. We topologize $\mathcal{D}^n(U)$ with the usual locally convex topology, characterized by the assertion that $\omega_k = \sum_{\alpha \in I_{n,P}} \omega_{k\alpha} dx^\alpha \to \omega = \sum_{\alpha \in I_{n,P}} \omega_\alpha dx^\alpha$ if there is a fixed compact $K \subset U$ such that $\operatorname{spt} \omega_{k\alpha} \subset K \ \forall \alpha \in I_{n,P}$, $k \geq 1$, and if $\lim D^\beta \omega_{k\alpha} = D^\beta \omega_\alpha$ uniformly in $K \ \forall \alpha \in I_{n,P}$ and every multi-index β. For any $\omega \in \mathcal{D}^n(U)$, we define

1.15 $$|\omega| = \sup_{x \in U} \sqrt{\omega(x) \cdot \omega(x)}$$

If $f : U \to V$ is smooth (U, V open in $\mathbb{R}^P$, $\mathbb{R}^Q$ respectively) and if f is proper (i.e. $f^{-1}(K)$ is a compact subset of U whenever K is a compact subset of V) then $f^\# \omega \in \mathcal{D}^n(U)$ whenever $\omega \in \mathcal{D}^n(V)$.

2 General Currents

Throughout this section U is an open subset of $\mathbb{R}^P$.

2.1 Definition: An n-dimensional current (briefly called an n-current) in U is a continuous linear functional on $\mathcal{D}^n(U)$. The set of such n-currents (i.e. the dual space of $\mathcal{D}^n(U)$) will be denoted $\mathcal{D}_n(U)$.

Note that in case $n = 0$ the n-currents in U are just the Schwartz distributions on U. More importantly though, the n-currents, $n \geq 1$, can be interpreted as a generalization of the n-dimensional oriented submanifolds M having locally finite $\mathcal{H}^n$-measure in U. Indeed given such an $M \subset U$ with orientation ξ (thus $\xi(x)$ is continuous on M with $\xi(x) = \pm \tau_1 \wedge \cdots \wedge \tau_n \ \forall x \in M$, where $\tau_1, \ldots, \tau_n$ is an orthonormal basis for $T_x M$)[1], there is a corresponding n-current $[\![M]\!] \in \mathcal{D}_n(U)$ defined by

2.2 $$[\![M]\!](\omega) = \int_M \langle \omega(x), \xi(x) \rangle \, d\mathcal{H}^n(x), \quad \omega \in \mathcal{D}^n(U),$$

where $\langle\, , \rangle$ denotes the dual pairing for $\Lambda^n(\mathbb{R}^P)$, $\Lambda_n(\mathbb{R}^P)$ as in 1.7. (That is, the n-current $[\![M]\!]$ is obtained by integration of n-forms over M in the usual sense of differential geometry: $[\![M]\!](\omega) = \int_M \omega$ in the usual notation of differential geometry.)

Motivated by the classical Stokes' theorem ($\int_M d\omega = \int_{\partial M} \omega$ if M is a compact smooth manifold with smooth boundary) we are led (by 2.2) to quite generally define the boundary ∂T of an n-current $T \in \mathcal{D}_n(U)$ by

2.3 $$\partial T(\omega) = T(d\omega), \quad \omega \in \mathcal{D}^n(U)$$

(and $\partial T = 0$ if $n = 0$); thus $\partial T \in \mathcal{D}_{n-1}(U)$ if $T \in \mathcal{D}_n(U)$. Here and subsequently we define $\mathcal{D}_{n-1}(U) = 0$ in case $n = 0$.

Notice that $\partial^2 T = 0$ by 1.12.

Again motivated by the special example $T = [\![M]\!]$ as in 2.2 we define the *mass* of T, $\mathbb{M}(T)$, for $T \in \mathcal{D}_n(U)$ by

2.4 $$\mathbb{M}(T) = \sup_{|\omega| \leq 1,\, \omega \in \mathcal{D}^n(U)} T(\omega)$$

(so that $\mathbb{M}(T) = \mathcal{H}^n(M)$ in case $T = [\![M]\!]$ as in 2.2). More generally for any open $W \subset U$ we define

2.5 $$\mathbb{M}_W(T) = \sup_{|\omega| \leq 1,\, \omega \in \mathcal{D}^n(U), \operatorname{spt} \omega \subset W} T(\omega)$$

[1] Thus $\xi(x) \in \Lambda_n(T_x M)$; notice this differs from the usual convention of differential geometry where we would take $\xi(x) \in \Lambda^n(T_x M)$.

2.6 Remark: We here adopt the definition of $\mathbb{M}(T)$ using the inner product norm $|\omega|$, but notice that there is some flexibility in this; we would still get the "correct" value $\mathcal{H}^n(M)$ for the case $T = [\![M]\!]$ if we were to make the definition $\mathbb{M}(T) = \sup_{\|\omega(x)\|\le 1 \omega\in\mathcal{D}^n(U)} T(\omega)$, where $\|\omega(x)\|$ denotes the comass norm of ω at x; thus

$$\|\omega\| = \sup_{\xi\in\Lambda_n(\mathbb{R}^P),\,|\xi|=1,\,\xi\text{ simple}} \langle\omega,\xi\rangle.$$

Indeed in general this works (for $T = [\![M]\!]$) provided only that $\|\ \|$ is a norm for $\Lambda^n(\mathbb{R}^P)$ with the properties:

(a) $\langle\omega,\xi\rangle \le \|\omega\|\,|\xi|$ whenever $\xi \in \Lambda_n(\mathbb{R}^P)$ is *simple*

(b) For each fixed simple $\xi \in \Lambda_n(\mathbb{R}^P)$, equality holds in (a) for some $\omega \neq 0$.

Evidently the inner product norm and the comass norm are two such norms, but the comass norm is the *smallest* possible norm for $\Lambda^n(\mathbb{R}^P)$ having these properties, which gives *maximality* of the corresponding definition of $\mathbb{M}(T)$. The reader is warned that $\mathbb{M}(T)$ is usually defined in terms of the comass norm—this makes no significant difference to later discussion here but of course there will be contexts in which the difference becomes significant.

Notice that by the Riesz Representation Theorem 4.14 of Ch. 1 we have that if $T \in \mathcal{D}_n(U)$ satisfies $\mathbb{M}_W(T) < \infty$ for every open $W \subset\subset U$, then there is a Radon measure μ_T on U and μ_T-measurable function $\vec{T}$ with values in $\Lambda_n(\mathbb{R}^P)$, $|\vec{T}| = 1$ μ_T-a.e., such that

2.7 $$T(\omega) = \int_U \langle\omega(x), \vec{T}(x)\rangle\, d\mu_T(x).$$

μ_T is characterized by

2.8 $$\mu_T(W) = \mathbb{M}_W(T)\ \Big(= \sup_{\omega\in\mathcal{D}^n(U),\,\|\omega\|\le 1,\,\operatorname{spt}\omega\subset W} T(\omega)\Big)$$

for any open W with $\overline{W}$ a compact subset of U. In particular

$$\mu_T(U) = \mathbb{M}(T).$$

Notice that for such a T we can define, for any μ_T-measurable subset A of U (and in particular for any Borel set $A \subset U$), a new current $T \llcorner A \in \mathcal{D}_n(U)$ by

2.9 $$(T \llcorner A)(\omega) = \int_A \langle\omega, \vec{T}\rangle\, d\mu_T.$$

More generally, if φ is any locally μ_T-integrable function on U then we can define $T \llcorner \varphi \in \mathcal{D}_n(U)$ by

2.10 $$(T \llcorner \varphi)(\omega) = \int \langle\omega, \xi\rangle\varphi\, d\mu_T.$$

Given $T \in \mathcal{D}_n(U)$ we define the *support*, $\operatorname{spt} T$, of T to be the relatively closed subset of U defined by

2.11 $$\operatorname{spt} T = U \setminus \cup W$$

where the union is over all open sets $W \subset\subset U$ such that $T(\omega) = 0$ whenever $\omega \in \mathcal{D}^n(U)$ with $\operatorname{spt}\omega \subset W$. Notice that if $\mathbb{M}_W(T) < \infty$ for each $W \subset\subset U$ and if μ_T is the corresponding total variation measure (as in 2.7, 2.8) then

2.12 $$\operatorname{spt} T = \operatorname{spt}\mu_T$$

where $\operatorname{spt}\mu_T$ is the support of μ_T in the usual sense of Radon measures in U.

Given a sequence $\{T_q\} \subset \mathcal{D}_n(U)$, we write $T_q \rightharpoonup T$ in U ($T \in \mathcal{D}_n(U)$) if $\{T_q\}$ converges weakly to T in the usual sense of distributions:

2.13 $$T_q \rightharpoonup T \iff \lim T_q(\omega) = T(\omega)\ \forall\omega \in \mathcal{D}^n(U).$$

Notice that mass is trivially lower semi-continuous with respect to weak convergence: if $T_q \rightharpoonup T$ in U then

2.14 $$\mathbb{M}_W(T) \le \liminf_{q\to\infty} M_W(T_q)\ \forall \text{ open } W \subset U.$$

We also observe that if $\sup_q \mathbb{M}_W(T_q) < \infty$ for each open $W \subset\subset U$ then by 2.7 the distribution convergence 2.13 is equivalent to weak* convergence with respect to continuous forms with compact support (i.e. $T_q(\omega) \to T(\omega)$ for all continuous n-forms ω on U with compact support), and hence by applying the standard Banach-Alaoglu theorem [Roy88] (in the Banach spaces $\mathcal{M}_n(W) = \{T \in \mathcal{D}_n(W) : \mathbb{M}_W(T) < \infty\}$, $W \subset\subset U$) we deduce

2.15 Lemma. *If $\{T_q\} \subset \mathcal{D}_n(U)$ and $\sup_{q\ge 1}\mathbb{M}_W(T_q) < \infty$ for each $W \subset\subset U$, then there is a subsequence $\{T_{q'}\}$ and a $T \in \mathcal{D}_n(U)$ such that*

$$\int_U \langle\omega, \vec{T}_{q'}\rangle\, d\mu_{T_{q'}} \to \int_U \langle\omega, \vec{T}\rangle\, d\mu_T$$

for each continuous n-form ω with compact support in U.

The following terminology will be used frequently:

2.16 Terminology: Given $T_1 \in \mathcal{D}_n(U_1)$, $T_2 \in \mathcal{D}_n(U_2)$ and an open $W \subset U_1 \cap U_2$, we say $T_1 = T_2$ in W if $T_1(\omega) = T_2(\omega)$ whenever ω is a smooth n-form in $\mathbb{R}^{n+\ell}$ with $\operatorname{spt}\omega \subset W$.

Next we want to describe the cartesian product of currents $T_1 \in \mathcal{D}_s(U_1)$, $T_2 \in \mathcal{D}_t(U_2)$, $U_1 \subset \mathbb{R}^{P_1}$, $U_2 \subset \mathbb{R}^{P_2}$ open. We are motivated by the case when $T_1 = [\![M_1]\!]$

and $T_2 = [\![M_2]\!]$ (Cf. 2.2) where M_1, M_2 are oriented submanifolds of dimension s, t respectively. We want to define $T_1 \times T_2 \in \mathcal{D}_{s+t}(U_1 \times U_2)$ in such a way that for this special case (when $T_j = [\![M_j]\!]$) we get $[\![M_1]\!] \times [\![M_2]\!] = [\![M_1 \times M_2]\!]$. Since $M_1 \times M_2$ has the natural orienting $(s+t)$-vector $p_\#(\xi) \wedge q_\#(\eta)$, where ξ and η are the orienting s-vector and t-vector for M_1, M_2 respectively, and where $p(x) = (x, 0)$, $x \in \mathbb{R}^{P_1}$, and $q(y) = (0, y)$, $y \in \mathbb{R}^{P_2}$, we are thus led to the following definition:

2.17 Definition: If $\omega \in \mathcal{D}^{s+t}(U_1 \times U_2)$ is written in the form

$$\omega = \textstyle\sum_{(\alpha,\beta) \in I_{s',P_1} \times I_{t',P_2}, s'+t'=s+t} a_{\alpha\beta}(x, y)\, dx^\alpha \wedge dy^\beta$$

then we define

$$S \times T(\omega) = T\big(\textstyle\sum_{\beta \in I_{t,P_2}} S\big(\sum_{\alpha \in I_{s,P_1}} a_{\alpha\beta}(x, y) dx^\alpha\big) dy^\beta\big),$$

which makes sense because if $\operatorname{spt}\omega = K$ then $K \subset P(K) \times Q(K)$, where P denotes the projection $(x, y) \mapsto x$ of $U_1 \times U_2 \to U_1$ and Q denotes the projection $(x, y) \mapsto y$ of $U_1 \times U_2 \to U_2$, and one can check that $S(\sum_{\alpha \in I_{s,P_1}} \omega_{\alpha\beta}(x, y) dx^\alpha)$ is a $C_c^\infty(U_2)$ function of y with support in $Q(K)$.

Notice in particular this gives, for $\omega_1 \in \mathcal{D}^{s'}(U_1)$, $\omega_2 \in \mathcal{D}^{t'}(U_2)$ with $s' + t' = s + t$ and with P, Q as above,

2.18 $$S \times T((P^\# \omega_1) \wedge (Q^\# \omega_2)) = \begin{cases} S(\omega_1) T(\omega_2) & \text{if } (s', t') = (s, t) \\ 0 & \text{if } (s', t') \neq (s, t). \end{cases}$$

One readily checks, using Definition 2.17 and the definition of ∂ (in 2.3), that

2.19 $$\partial(S \times T) = (\partial S) \times T + (-1)^s S \times \partial T.$$

(Notice this is valid also in case r or $s = 0$ if we interpret the appropriate terms as zero; e.g. if $s = 0$ then $\partial(S \times T) = S \times \partial T$.) Also, by 2.17 and 2.18,

2.20 $$\mathbb{M}_{W_1 \times W_2}(S \times T) = \mathbb{M}_{W_1}(S) \mathbb{M}_{W_2}(T)$$

for any open $W_1 \subset\subset U_1$, $W_2 \subset\subset U_2$.

An important special case of 2.19 occurs when we take $T \in \mathcal{D}_n(U)$, $U \subset \mathbb{R}^P$, and we let $[\![(0, 1)]\!]$ be the 1-current defined as in 2.3 with $M = (0, 1) \subset \mathbb{R}$ ($(0, 1)$ having its usual orientation). Then 2.19 gives

2.21 $$\begin{aligned} \partial([\![(0, 1)]\!] \times T) &= (\{1\} - \{0\}) \times T - [\![(0, 1)]\!] \times \partial T \\ &\equiv \{1\} \times T - \{0\} \times T - [\![(0, 1)]\!] \times \partial T. \end{aligned}$$

Here and subsequently $\{p\}$, for a point $p \in U$, means the 0-current $\in \mathcal{D}_0(U)$ defined by

$$\{p\}(\omega) = \omega(p), \quad \omega \in \mathcal{D}^0(U) \ \big(= C_c^\infty(U)\big).$$

Then if $\omega = \sum_{\alpha \in I_{s,P}, \beta \in I_{t,Q}, s+t=n} \omega_{\alpha\beta}(x, y) dx^\alpha dy^\beta \in \mathcal{D}^n(U \times V)$ with $U \subset \mathbb{R}^P$, $V \subset \mathbb{R}^Q$ open, and if $T \in \mathcal{D}_n(V)$, then by Definition 2.17 in the case $s = 0$ we have

2.22 $$(\{p\} \times T)(\omega) = T\big(\textstyle\sum_{\beta \in I_{n,Q}} \omega_{\alpha,\beta}(p, y) dy^\beta\big).$$

Next we want to discuss the notion of "pushing forward" a current T via a smooth map $f : U \to V$, $U \subset \mathbb{R}^P$, $V \subset \mathbb{R}^Q$ open. The main restriction needed is that $f \,|\, \operatorname{spt} T$ is *proper*; that is $f^{-1}(K) \cap \operatorname{spt} T$ is a compact subset of U whenever K is a compact subset of V. Assuming this, we can define

2.23 $$f_\# T(\omega) = T\big(\zeta f^\# \omega\big) \ \forall \omega \in \mathcal{D}^n(V),$$

where ζ is any function $\in C_c^\infty(U)$ such that ζ is identically equal to one in a neighborhood of the compact set $\operatorname{spt} T \cap \operatorname{spt} f^\# \omega$. One easily checks that the definition of $f_\# T$ in 2.23 is independent of the particular choice of ζ.

2.24 Remarks: (1) Notice that $\partial f_\# T = f_\# \partial T$ whenever f, T are as in 2.23.

(2) If $\mathbb{M}_W(T) < \infty$ for each $W \subset\subset U$, so that T has a representation as in 2.7, then it is straightforward to check that $f_\# T$ is given explicitly by

$$\begin{aligned} f_\# T(\omega) &= \int \langle f^\# \omega, \vec{T} \rangle \, d\mu_T \\ &= \int \langle \omega|_{f(x)}, df_{x\#} \vec{T}(x) \rangle \, d\mu_T(x). \end{aligned}$$

Thus if $\mathbb{M}_W(T) < \infty \, \forall W \subset\subset U$ we can make sense of $f_\# T$ in case f is merely C^1 with $f \,|\, \operatorname{spt} T$ proper.

(3) If $T = [\![M]\!]$ as in 2.2, the above remark (2) tells us that if $f \,|\, \overline{M} \cap U$ is proper, then

$$f_\# T(\omega) = \int_M \langle \omega|_{f(x)}, df_{x\#} \xi(x) \rangle \, d\mathcal{H}^n(x),$$

where ξ is the orientation for M. Notice that this makes sense if f is only Lipschitz (by virtue of Rademacher's Theorem 1.4 of Ch. 2). If f is 1:1 and if J_f is the Jacobian of f as in 3.3 of Ch. 2, then the area formula evidently tells us that (since $df_{x\#}\xi(x) = J_f(x) \eta(f(x))$, where η is the orientation for $f(M_+)$, $M_+ = \{x \in M : J_f(x) > 0\}$, induced by f)

$$f_\# T(\omega) = \int_{f(M_+)} \langle \omega(y), \eta(y) \rangle \, d\mathcal{H}^n(y).$$

(Which confirms that our definition of $f_\# T$ is "correct.")

We can now derive the important homotopy formula for currents as follows: If $f, g : U \to V$ are smooth ($V \subset \mathbb{R}^Q$) and $h : [0,1] \times U \to V$ is smooth with $h(0,x) \equiv f(x)$, $h(1,x) \equiv g(x)$, if $T \in \mathcal{D}_n(U)$, and if $h|[0,1] \times \operatorname{spt} T$ is proper, then, by 2.21 and 2.22,

$$\begin{aligned}\partial h_\#([\![(0,1)]\!] \times T) &= h_\# \partial([\![(0,1)]\!] \times T)\\ &= h_\#(\{1\} \times T - \{0\} \times T - [\![(0,1)]\!] \times \partial T)\\ &\equiv g_\# T - f_\# T - h_\#([\![(0,1)]\!] \times \partial T).\end{aligned}$$

Thus we obtain the *homotopy formula*

2.25 $$g_\# T - f_\# T = \partial h_\#([\![(0,1)]\!] \times T) + h_\#([\![(0,1)]\!] \times \partial T).$$

Notice that an important case of the above is given by

2.26 $$h(t,x) = tg(x) + (1-t)f(x) = f(x) + t(g(x) - f(x))$$

(i.e. h is an "affine homotopy" from f to g). In this case we note that if $h|\operatorname{spt} T$ is a proper map into V then $W \subset\subset V \Rightarrow \operatorname{spt}([\![(0,1)]\!] \times T) \cap h^{-1}(W) \subset\subset [0,1] \times U$ and hence $\operatorname{spt} T \cap p(h^{-1}(W)) \subset\subset U$, where p is the projection $(t,x) \mapsto x$. Then by the integral representation 2.7 and Remark 2.24 (2) above we have, for any open $W \subset\subset V$,

2.27 $$\mathbb{M}_W(h_\#[\![(0,1)]\!] \times T) \le \sup_{x \in \operatorname{spt} T \cap W_h} |f - g| \cdot \sup_{x \in \operatorname{spt} T \cap W_h} (|df_x| + |dg_x|)^n \,\mathbb{M}_{W_h}(T),$$

where $W_h = p(h^{-1}(W))$, with $p : (t,x) \mapsto x$. Indeed $\overrightarrow{[\![(0,1)]\!] \times T} = e_1 \wedge \vec{T}$ and $\mu_{[\![(0,1)]\!] \times T} = \mathcal{L}^1 \times \mu_T$, so by 2.24 (1) we have, for any $\omega \in \mathcal{D}^n(V)$,

$$\begin{aligned}h_\#([\![(0,1)]\!] \times T)(\omega) &= \int_0^1 \int_U \langle \omega_{|h(t,x)}, df_{(t,x)\#}(e_1 \wedge \vec{T}(x)) \rangle \, d\mu_T(x)dt\\ &= \int_0^1 \int_U \langle \omega_{|h(t,x)}, (g(x) - f(x))\\ &\qquad \wedge (t dg_x + (1-t)df_x)_\# \vec{T}(x) \rangle \, d\mu_T(x)dt,\end{aligned}$$

and 2.27 follows immediately.

We now give a couple of important applications of the above homotopy formula.

2.28 Lemma: *If $T \in \mathcal{D}_n(U)$, $\mathbb{M}_W(T)$, $\mathbb{M}_W(\partial T) < \infty$ $\forall W \subset\subset U$ and if $f, g : U \to V$ are C^1 and if h is as in 2.26 with $h(\operatorname{spt} T) \subset V$ and $h|\operatorname{spt} T$ proper, then $f_\# T = g_\# T$. (Note that $f_\# T$, $g_\# T$ are well-defined by 2.24 (2).)*

Proof: By the homotopy formula 2.25 we have, with $h(t,x) = tg(x) + (1-t)f(x)$,

$$\begin{aligned}g_\# T(\omega) - f_\# T(\omega) &= \partial h_\#([\![(0,1)]\!] \times T) + h_\#([\![(0,1)]\!] \times \partial T)(\omega)\\ &= h_\#([\![(0,1)]\!] \times T)(d\omega) + h_\#([\![(0,1)]\!] \times \partial T)(\omega),\end{aligned}$$

so that, by 2.27,

$$|f_\# T(\omega) - g_\# T(\omega)| \le c\big(\mathbb{M}(T)|d\omega| + \mathbb{M}(\partial T)|\omega|\big) \sup_{x \in \operatorname{spt} T} |f - g| = 0$$

since $f = g$ on $\operatorname{spt} T$. □

The homotopy formula also enables us to define $f_\# T$ in case f is merely Lipschitz, provided $f|\operatorname{spt} T$ is proper and $\mathbb{M}_W(T)$, $\mathbb{M}_W(\partial T) < \infty$ $\forall W \subset\subset U$. In the following lemma we let $f_\sigma = f * \varphi^{(\sigma)}$, $\varphi^{(\sigma)}(x) = \sigma^{-P}\varphi(\sigma^{-1}x)$, with φ a mollifier as in §2 of Ch. 2.

2.29 Lemma. *If $T \in \mathcal{D}_n(U)$, $\mathbb{M}_W(T)$, $\mathbb{M}_W(\partial T) < \infty$ $\forall W \subset\subset U$, and if $f : U \to V$ is Lipschitz with $f|\operatorname{spt} T$ proper, then $\lim_{\sigma \downarrow 0} f_{\sigma\#} T(\omega)$ exists for each $\omega \in \mathcal{D}^n(V)$; $f_\# T(\omega)$ is defined to be this limit; then $\operatorname{spt} f_\# T \subset f\,(\operatorname{spt} T)$ and $\mathbb{M}_W(f_\# T) \le \big(\operatorname{ess\,sup}_{f^{-1}(W)} |Df|\big)^n \mathbb{M}_{f^{-1}(W)}(T)$ $\forall W \subset\subset V$.*

Proof: If σ, τ are sufficiently small (depending on ω) then the homotopy formula gives

$$f_{\sigma\#} T(\omega) - f_{\tau\#} T(\omega) = h_\#([\![(0,1)]\!] \times T)(d\omega) + h_\#([\![(0,1)]\!] \times \partial T)(\omega)$$

where $h : [0,1] \times U \to V$ is defined by $h(t,x) = tf_\sigma(x) + (1-t)f_\tau(x)$. Then by 2.27, for sufficiently small σ, τ, we have

$$|f_{\sigma\#} T(\omega) - f_{\tau\#} T(\omega)| \le C \sup_{f^{-1}(K) \cap \operatorname{spt} T} |f_\sigma - f_\tau| \cdot (\operatorname{Lip} f)^n,$$

where K is a compact subset of V with $\operatorname{spt} \omega \subset \operatorname{interior}(K)$. Since $f_\sigma \to f$ uniformly on compact subsets of U, the result now clearly follows. □

Next we want to define the notion of the *cone* over a given current $T \in \mathcal{D}_n(U)$. We want to define this in such a way that if $T = [\![M]\!]$ where M is a submanifold of $\mathbb{S}^{P-1} \subset \mathbb{R}^P$ then the cone over T is just $[\![C_M]\!]$, $C_M = \{\lambda x : x \in M,\ 0 < \lambda \le 1\}$. We are thus led generally to make the definition that the cone over T, denoted $0 \mathbin{⨳} T$, is defined by

2.30 $$0 \mathbin{⨳} T = h_\#([\![(0,1)]\!] \times T)$$

whenever $T \in \mathcal{D}_n(U)$ with U star-shaped relative to 0 and $\operatorname{spt} T$ compact, where $h : [0,1] \times \mathbb{R}^P \to \mathbb{R}^P$ is defined by $h(t,x) = tx$. Notice that h is an affine homotopy

$tg(x) + (1-t)f(x)$, where $g(x) = x$ and $f(x) = 0$. Thus $0 \times\!\!\!\!\times T \in \mathcal{D}_{n+1}(U)$ and (by the homotopy formula)

2.31 $$\partial(0 \times\!\!\!\!\times T) = T - 0 \times\!\!\!\!\times \partial T.$$

Notice in particular that, with $R = 0 \times\!\!\!\!\times T$, we have thus established that

2.32 U star-shaped relative to 0 and $T \in \mathcal{D}_n(U)$ with spt T compact and $\partial T = 0 \Rightarrow \exists R \in \mathcal{D}_{n+1}(U)$ with $\partial R = T$.

As a final application of the homotopy formula we have the following lemma which is useful for checking if a given current of locally finite mass is conical—i.e. invariant under homotheties $\eta_{0,\lambda}$:

2.33 Lemma. *Suppose $C \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$ with $M_{\breve{B}_R(0)} < \infty$ for each $R > 0$, $\partial C = 0$, and $x \wedge \vec{C}|_x = 0$ μ_C-a.e. Then $\eta_{0,\lambda\#}C = C$ for each $\lambda > 0$.*

Proof: We apply the homotopy formula 2.25 with $f(x) = x$ and $g(x) = \lambda^{-1}x$, and $h(t,x) = tg(x) + (1-t)f(x)$. Then

$$\eta_{0,\lambda\#}C - C = \partial h_\#([\![(0,1)]\!] \times C).$$

The right side here is zero because $\overrightarrow{[\![(0,1)]\!] \times C} = e_1 \wedge \vec{C}$, and hence

$$h_\#|_{(t,x)}\, \overrightarrow{[\![(0,1)]\!] \times C}|_{t,x} = (1 + t(\lambda^{-1} - 1))^n(\lambda^{-1} - 1)\, x \wedge \vec{C}|_x = 0. \ \square$$

The following Constancy Theorem is very useful:

2.34 Theorem. *If U is open in $\mathbb{R}^n$ (i.e. $P = n$), if U is connected, if $T \in \mathcal{D}_n(U)$ and $\partial T = 0$, then there is a constant c such that $T = c[\![U]\!]$ (using the notation of 2.2 in the special case $P = n$, $M = U$; U is of course equipped with the standard orientation $e_1 \wedge \cdots \wedge e_n$).*

Proof: Let $\varphi^{(\sigma)}(x) = \sigma^{-n}\varphi(\sigma^{-1}x)$, with φ a mollifier as in §2 of Ch. 2. For any ball $B_\rho(x_0) \subset U$ first pick $R > \rho$ with $B_R(x_0) \subset U$ and take $a \in L^1(\mathbb{R}^n)$ with $a \equiv 0$ on $\mathbb{R}^n \setminus B_\rho(x_0)$. Then we have $a_\sigma \in C_c^\infty(U)$ for $\sigma < R - \rho$) ($a_\sigma = \varphi^{(\sigma)} * a$), and $D^\beta a_\sigma = (D^\beta \varphi^{(\sigma)}) * a$ for each multi-index β, so if $a_j \to a$ in $L^1(\breve{B}_\rho(x_0))$ with $a_j \equiv 0$ on $\mathbb{R}^n \setminus \breve{B}_\rho(x_0)$ then $a_{j\sigma}dx^1 \wedge \cdots \wedge dx^n \to a_\sigma dx^1 \wedge \cdots \wedge dx^n$ in $\mathcal{D}^n(U)$, and hence $T(a_{j\sigma}dx^1 \wedge \cdots \wedge dx^n) \to T(a_\sigma dx^1 \wedge \cdots \wedge dx^n)$. Thus the functional $F_\sigma : L^1(\breve{B}_\rho(x_0)) \to \mathbb{R}$ defined by $F_\sigma(a) = T(a_\sigma dx^1 \wedge \cdots \wedge dx^n)$ is a bounded linear functional on $L^1(\breve{B}_\rho(x_0))$ for $\sigma < R - \rho$, and by the Riesz representation theorem for $L^1(\breve{B}_\rho(x_0))$ there is a bounded measurable function $\theta^{(\sigma)}$ in $\breve{B}_\rho(x_0)$ with $F_\sigma(a) = \int_{\breve{B}_\rho(x_0)} a\theta^{(\sigma)}\, d\mathcal{L}^n$ for $a \in L^1(\breve{B}_\rho(x_0))$, and hence

(1) $$T(a_\sigma dx^1 \wedge \cdots \wedge dx^n) = \int a\theta^{(\sigma)}\, d\mathcal{L}^n,\ a \in C_c^\infty(W).$$

Now let $a \in C_c^\infty(\breve{B}_\rho(x_0))$ and for $j = 1, \ldots, n$ let $\omega_{j\sigma} = (-1)^{j-1}a_\sigma dx^1 \wedge \cdots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \cdots \wedge dx^n$, and observe that $d\omega_{j\sigma} = (D_j a)_\sigma dx^1 \wedge \cdots \wedge \wedge dx^n$, so (1) with $D_j a$ in place of a implies

(2) $$\int D_j a\, \theta^{(\sigma)}\, d\mathcal{L}^n = T(d\omega_{j\sigma}) = \partial T(\omega_{j\sigma}) = 0, \quad j = 1, \ldots, n.$$

If $\tau < R - \rho$ and $x \in \breve{B}_\rho(x_0)$, then $a_\tau = \varphi^{(\tau)} * a *$ is a $C_c^\infty(B_\rho(x_0))$ function of x and hence a legitimate choice for a in (2), so in fact in that case (2) says $D_j(\theta^{(\sigma)})_\tau = 0$ on $\breve{B}_\rho(x_0)$ hence $(\theta^{(\sigma)})_\tau$ is constant on $\breve{B}_\rho(x_0)$, hence (letting $\tau \downarrow 0$) we see that $\theta^{(\sigma)}$ is constant c_σ on $\breve{B}_\rho(x_0)$. Letting $\sigma \downarrow 0$ we deduce that $T = c_{\rho,x_0}[\![\breve{B}_\rho(x_0)]\!]$ (c_{ρ,x_0} constant) on each ball $\breve{B}_\rho(x_0)$ with $B_\rho(x_0) \subset U$, and the result follows. $\square$

2.35 Remark: Notice that if we merely have $\mathbb{M}_W(\partial T) < \infty$ for each $W \subset\subset U$ then the obvious modifications of the above argument give first that

$$\left|\int D_j a\, \theta_\sigma\, d\mathcal{L}^n\right| \le C \sup|a|\, \mathbb{M}(\partial T)$$

with C independent of σ, for $a \in C_c^\infty(U)$ such that $\operatorname{dist}(\operatorname{spt} a, \partial U) > \sigma$. Next we must justify that θ_{σ_k} is bounded in $\mathcal{L}^1(B)$ for each open ball $B \subset\subset U$. Indeed by 2.7 of Ch. 2 and $\mathbb{M}_B(\partial T) < \infty$, there are constants λ_k such that $\theta_{\sigma_k} - \lambda_k$ is bounded in $\mathcal{L}^1(B)$, and hence $T_{\sigma_k} - \lambda_k[\![B]\!]$ has bounded mass in B. But $T_{\sigma_k} \rightharpoonup T$ and hence $\{\lambda_k\}$ is bounded. Thus (see §2 of Ch. 2 and in particular 2.6 of Ch. 2) we deduce that $\theta_{\sigma_k} \to \theta$ in $L^1_{\text{loc}}(U)$ (for some sequence $\sigma_k \downarrow 0$), with $\theta \in BV_{\text{loc}}(U)$, and

(‡) $$T(\omega) = \int a\, \theta\, d\mathcal{L}^n, \quad \omega = a\, dx^1 \wedge \cdots \wedge dx^n \in \mathcal{D}^n(U).$$

Using the definition of $\mathbb{M}(\partial T)$, we easily then check that $\mathbb{M}_W(\partial T) = |D\theta|(W)$ for each open $W \subset U$ (and $\mathbb{M}_W(T) = \int_W |\theta|\, d\mathcal{L}^n$). Indeed in the present case $n = P$, any $\omega \in \mathcal{D}^{n-1}(U)$ can be written $\omega = \sum_{j=1}^n (-1)^j a_j\, dx^1 \wedge \cdots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \cdots \wedge dx^n$ for suitable $a_j \in C_c^\infty(U)$, and $d\omega = \operatorname{div} \underline{a}\, dx^1 \wedge \cdots \wedge dx^n$ for such ω ($\underline{a} = (a_1, \ldots, a_n)$). Therefore by (‡) above we have

$$\partial T(\omega) = T(d\omega) = \int \operatorname{div} \underline{a}\, \theta\, d\mathcal{L}^n$$

and the assertion $\mathbb{M}_W(\partial T) = |D\theta|(W)$ then follows directly from the definition of $\mathbb{M}_W(\partial T)$ and $|D\theta|$ (in §2 of Ch. 2).

In the following lemma, for $\alpha = (i_1, \ldots, i_n) \in \mathbb{Z}^n$ with $1 \le i_1 < i_2 < \cdots < i_n \le P$, we let p_α denote the orthogonal projection of $\mathbb{R}^P$ onto $\mathbb{R}^n$ given by

$$(x^1, \ldots, x^P) \mapsto (x^{i_1}, \ldots, x^{i_n}).$$

2.36 Lemma. *Suppose E is a closed subset of U, U open in $\mathbb{R}^P$, with $\mathcal{L}^n(p_\alpha(E)) = 0$ for each multi-index $\alpha = (i_1, \ldots, i_n)$, $1 \le i_1 < i_2 < \cdots < i_n \le P$. Then $T \llcorner E = 0$ whenever $T \in \mathcal{D}_n(U)$ with $\mathbb{M}_W(T), \mathbb{M}_W(\partial T) < \infty$ for every $W \subset\subset U$.*

2.37 Remarks: (1) The hypothesis $\mathcal{L}^n(p_\alpha(E)) = 0$ is trivially satisfied if $\mathcal{H}^n(E) = 0$, so in particular we deduce $T \llcorner E = 0$ if $T \in \mathcal{D}_n(U)$ with $\mathbb{M}_W(T), \mathbb{M}_W(\partial T) < \infty$ $\forall W \subset\subset U$ and $\mathcal{H}^n(E) = 0$. Since μ_T is Borel regular and finite on compact subsets of U, 1.15 of Ch.1 implies that $\mu_T(C) = \sup_{E \text{ closed } E \subset C} \mu_T(E)$, hence the lemma implies that if C is a Borel set with $\mathcal{H}^n(C) = 0$ then $\mu_T(C) = 0$. That is, μ_T is absolutely continuous with respect to $\mathcal{H}^n$ in U provided $\mathbb{M}_W(T), \mathbb{M}_W(\partial T) < \infty$ $\forall W \subset\subset U$.

(2) Let Q be any orthogonal transformation of $\mathbb{R}^P$. Since $T \in \mathcal{D}_n(U) \Rightarrow Q_\# T \in \mathcal{D}_n(QU)$ and $\mathbb{M}_W(T) = \mathbb{M}_{QW}(Q_\# T)$ for each $W \subset U$. So if $\mathbb{M}_W(T) < \infty$ for each $W \subset\subset U$ we have $\mu_{Q_\# T}(Q(A)) = \mu_T(A)$ for each $A \subset U$, hence the above lemma guarantees $\mathcal{L}^n(Q(E)) = 0$ for each $\alpha \Rightarrow \mu_T(E) = 0$

Proof of 2.36: Let $\omega \in \mathcal{D}^n(U)$. Then we can write $\omega = \sum_{a \in I_{n,P}} \omega_\alpha dx^\alpha$ $\omega_\alpha \in C_c^\infty(U)$, so that

$$T(\omega) = \textstyle\sum_\alpha T(\omega_\alpha dx^\alpha) = \sum_\alpha (T \llcorner \omega_\alpha)(dx^\alpha)$$
$$= \textstyle\sum_\alpha (T \llcorner \omega_\alpha)\, p_\alpha^\# \, dy.$$

($dy = dy^1 \wedge \cdots \wedge dy^n$, $y^1, \ldots, y^n$ the standard coordinate functions in $\mathbb{R}^n$.) Thus

(1) $$T(\omega) = \textstyle\sum_\alpha p_{\alpha\#}(T \llcorner \omega_\alpha)(dy)$$

(which makes sense because $\operatorname{spt} T \llcorner \omega_\alpha \subset \operatorname{spt} \omega_\alpha =$ a compact subset of U). On the other hand

$$\mathbb{M}(\partial p_{\alpha\#}(T \llcorner \omega_\alpha)) = \mathbb{M}(p_{\alpha\#}\partial(T \llcorner \omega_\alpha))$$
$$\le \mathbb{M}(\partial(T \llcorner \omega_\alpha)) < \infty,$$

because for any $\eta \in \mathcal{D}^{n-1}(U)$

$$\partial(T \llcorner \omega_\alpha)(\eta) = (T \llcorner \omega_\alpha)(d\eta)$$
$$= T(\omega_\alpha d\eta)$$
$$= T(d(\omega_\alpha \eta)) - T(d\omega_\alpha \wedge \eta)$$
$$= \partial T(\omega_\alpha \eta) - T(d\omega_\alpha \wedge \eta);$$

thus in fact

$$\mathbb{M}_W(\partial(T \llcorner \omega_\alpha)) \le \mathbb{M}_W(\partial T)|\omega_\alpha| + \mathbb{M}_W(T)|d\omega_\alpha|.$$

Therefore by 2.35 we have $\theta_\alpha \in BV(p_\alpha(U))$ (depending on both α and ω_α) such that

$$p_{\alpha\#}(T \llcorner \omega_\alpha)(\eta) = \textstyle\int_{p_\alpha(U)} \langle \eta, e_1 \wedge \cdots \wedge e_n \rangle\, \theta_\alpha \, d\mathcal{L}^n,$$

and hence $p_{\alpha\#}(T \llcorner \omega_\alpha) \llcorner p_\alpha(E) = 0$ because $\mathcal{L}^n(p_\alpha(E)) = 0$. Then, assuming without loss of generality that E is closed,

(2) $$\begin{aligned} \mathbb{M}(p_{\alpha\#}(T \llcorner \omega_\alpha)) &\le \mathbb{M}(p_{\alpha\#}(T \llcorner \omega_\alpha) \llcorner (\mathbb{R}^n \setminus p_\alpha(E))) \\ &= \mathbb{M}(p_{\alpha\#}((T \llcorner \omega_\alpha) \llcorner (\mathbb{R}^P \setminus p_\alpha^{-1} p_\alpha(E)))) \\ &\le \mathbb{M}((T \llcorner \omega_\alpha) \llcorner (\mathbb{R}^P \setminus p_\alpha^{-1} p_\alpha E)) \\ &\le \mathbb{M}_W(T \llcorner (\mathbb{R}^P \setminus p_\alpha^{-1} p_\alpha E)) \cdot |\omega_\alpha| \\ &\le \mathbb{M}_W(T \llcorner (\mathbb{R}^P \setminus E)) \cdot |\omega_\alpha| \end{aligned}$$

for any W such that $\operatorname{spt} \omega \subset W \subset U$.

Combining (1) and (2) we then have

(3) $$\mathbb{M}_W(T) \le C\,\mathbb{M}_W(T \llcorner (\mathbb{R}^P \setminus E))$$

so that in particular

$$\mathbb{M}_W(T \llcorner E) \le C\,\mathbb{M}_W(T \llcorner (\mathbb{R}^P \setminus E)).$$

which says

(4) $$\mu_T(W \cap E) \le C\,\mu_T(W \setminus E)$$

Letting K be an arbitrary compact subset of E, we can choose $\{W_q\}$ so that $W_q \subset\subset U$, $W_{q+1} \subset W_q$, $\cap_{q=1}^\infty W_q = K$; using (4) with $W = W_q$ then gives $\mathbb{M}(T \llcorner K) = 0$, i.e., $\mu_T(K) = 0$. Since

$$\mu_T(E) = \sup_{K \text{ compact}, K \subset E} \mu_T(K),$$

by 1.15 of Ch.1, we thus have $\mu_T(E) = 0$. □

3 Integer Multiplicity Rectifiable Currents

In this section we want to develop the theory of integer multiplicity currents $T \in \mathcal{D}_n(U)$, which, roughly speaking are those currents obtained by assigning (in a $\mathcal{H}^n$-measurable fashion) an orientation to the tangent spaces $T_x V$ of an integer multiplicity varifold V. (See Ch. 4 for terminology.)

These currents are precisely those called locally *locally rectifiable currents* by Federer and Fleming [FF60], [Fed69].

Throughout this section $n \ge 1$, $k \ge 1$ are integers and U is an open subset of $\mathbb{R}^{n+\ell}$.

3.1 Definition: If $T \in \mathcal{D}_n(U)$ we say that T is an integer multiplicity rectifiable n-current (briefly an integer multiplicity current if it can be expressed

$$(\ddagger) \qquad T(\omega) = \int_M \langle \omega(x), \xi(x)\rangle\, \theta(x)\, d\mathcal{H}^n(x), \quad \omega \in \mathcal{D}^n(U),$$

where M is an $\mathcal{H}^n$-measurable countably n-rectifiable subset of U, θ is a locally $\mathcal{H}^n$-integrable positive integer-valued function, and $\xi : M \to \Lambda_n(\mathbb{R}^{n+\ell})$ is a $\mathcal{H}^n$-measurable function such that for $\mathcal{H}^n$-a.e. point $x \in M$, $\xi(x)$ can be expressed in the form $\tau_1 \wedge \cdots \wedge \tau_n$, where $\tau_1, \ldots, \tau_n$ form an orthonormal basis for the approximate tangent space $T_x M$. (See Ch. 3 and Ch. 4.) Thus $\xi (= \vec{T})$ orients the approximate tangent spaces of M in an $\mathcal{H}^n$-measurable way. The function θ in 3.1 ($\ddagger$) is called the *multiplicity* and ξ is called the *orientation* for T. If T is as in 3.1 ($\ddagger$) we shall often write

$$T = \underline{\tau}(M, \theta, \xi).$$

In this case

$$V = \underline{v}(M, \theta)$$

will be referred to as the *integer multiplicity varifold associated with* T.

3.2 Remarks: (1) If $T_1, T_2 \in \mathcal{D}_n(U)$ are integer multiplicity, then so is $p_1 T_1 + p_2 T_2$, $p_1, p_2 \in \mathbb{Z}$.

(2) If $T_1 = \underline{\tau}(M_1, \theta_1, \xi_1) \in \mathcal{D}_r(U)$, $T_2 = \underline{\tau}(M_2, \theta_2, \xi_2) \in \mathcal{D}_s(W)$ ($W \subset \mathbb{R}^Q$ open), then $T_1 \times T_2 \in \mathcal{D}_{r+s}(U \times W)$ is also integer multiplicity, and in fact

$$T_1 \times T_2 = \underline{\tau}(M_1 \times M_2, \theta_1\theta_2, p_\#(\xi_1) \wedge q_\#(\xi_2)),$$

where $p(x) = (x, 0)$ and $q(y) = (0, y)$ and $(\theta_1\theta_2)(x, y) = \theta_1(x)\theta_2(y)$.

(3) If $T = \tau(M, \theta, \xi) \in \mathcal{D}_n(U)$ is an integer multiplicity current then

$$\mathbb{M}_W(T) = \int_M \theta\, d\mathcal{H}^n = \mathbb{M}_W(V)\ \forall \text{ open } W \subset U,$$

where $V = \underline{v}(M, \theta)$ is the rectifiable varifold associated with T.

Next we want to discuss pushing forward an integer multiplicity $T = \underline{\tau}(M, \theta, \xi) \in \mathcal{D}_n(U)$ ($M \subset U$) by a Lipschitz map $f : U \to W$ such that $f \,|\, \operatorname{spt} T$ is proper. First, if f is C^1, 1:1, $f \,|\, \operatorname{spt} T$ is proper, M is an embedded C^1 submanifold, ξ is any $\mathcal{H}^n$-measurable orientation for M, and θ is any $\mathcal{H}^n$-measurable positive integer valued function on M, then we have, by Remark 2.24 (3),

3.3
$$\begin{aligned} f_\# T(\omega) &= \int_M \langle f^\# \omega, \xi\rangle\, \theta(x)\, d\mathcal{H}^n \\ &= \int_M \langle df_x^\# \omega_{|f(x)}, \xi_{|x}\rangle \theta(x)\, d\mathcal{H}^n \\ &= \int_M \langle \omega_{|f(x)}, df_{x\#}\xi_{|x}\rangle \theta(x)\, d\mathcal{H}^n. \end{aligned}$$

Now $\xi_{|x} = \pm\tau_1 \wedge \cdots \wedge \tau_n$, where $\tau_1, \ldots, \tau_n$ is an orthonormal basis for the tangent space $T_x M$, so

3.4
$$\begin{aligned} df_{x\#}\xi_{|x} &= \pm df_{x\#}\tau_1 \wedge \cdots \wedge df_{x\#}\tau_n \\ &= \pm D_{\tau_1} f(x) \wedge \cdots \wedge D_{\tau_n} f(x) \end{aligned}$$

which $= 0$ at points $x \in M$ where $J_f^M(x) = 0$, because $J_f^M(x) = 0 \Leftrightarrow \operatorname{rank}(d^M f_{|x}) < n$. On the other hand at points where $J_f^M(x) > 0$ the rank is n and hence there is $\rho > 0$ such that $f | M \cap \breve{B}_\rho(x)$ is a diffeomorphism onto an n-dimensional embedded C^1 manifold N, and at the image point $y = f(x)$ we let $\eta_1, \ldots, \eta_n$ be an orthonormal basis for $T_y N$. Then, since $D_{\tau_i} f(x) \in T_y N$, we have $D f_{\tau_i} f(x) = \sum_{j=1}^n D f_{\tau_i} \cdot \eta_j \eta_j$, and so

$$D_{\tau_1} f(x) \wedge \cdots \wedge D_{\tau_n} f(x) = \det(D_{\tau_i} \cdot \eta_j)\eta_1 \wedge \cdots \wedge \eta_n.$$

On the other hand

$$\begin{aligned} J_f^M(x) &= \sqrt{\det(D_{\tau_i} f(x) \cdot D_{\tau_j} f(x))} = \sqrt{(\det(D_{\tau_i} f(x) \cdot \eta_j))^2} \\ &= |\det(D_{\tau_i} f(x) \cdot \eta_j)|. \end{aligned}$$

Thus we see that 3.4 implies, at points $x \in M$ where $J_f^M(x) \neq 0$,

3.5
$$df_{x\#}\xi_{|x} = J_f^M(x)\eta,$$

where η is an orienting n-vector for N (so $\eta = \pm\eta_1 \wedge \cdots \wedge \eta_n$). η is called *the orientation for N induced by f* at each point x where $J_f^M(x) \neq 0$.

Now suppose $f : U \to W$ is Lipschitz, $T = \underline{\tau}(M, \theta, \xi) \in \mathcal{D}_n(U)$ ($M \subset U$) is an integer multiplicity current, and $f \,|\, \operatorname{spt} T$ is proper, then we can define $f_\# T \in \mathcal{D}_n(W)$ by

$$f_\# T(\omega) = \int_M \left\langle \omega_{|f(x)}, d^M f_{x\#}\xi(x)\right\rangle \theta(x)\, d\mathcal{H}^n(x).$$

Since $\left|d^M f_{x\#}\xi(x)\right| = J^M f(x)$ (as in §2 of Ch. 3) by the area formula this can be written

3.6
$$f_\# T(\omega) = \int_{f(M)} \langle \omega_{|y}, \textstyle\sum_{x \in f^{-1}(y) \cap M_+} \theta(x) \dfrac{d^M f_{x\#}\xi(x)}{\left|d^M f_{x\#}\xi(x)\right|}\rangle\, d\mathcal{H}^n(y),$$

where $M_+ = \{x \in M : J_M f(x) > 0\}$. Furthermore at points y where the approximate tangent space $T_y(f(M))$ exists (which is $\mathcal{H}^n$-a.e. y by virtue of the fact that $f(M)$ is countably n-rectifiable) and where $T_x M$, $d^M f_x$ exist $\forall x \in f^{-1}(y)$ (which is again for $\mathcal{H}^n$-a.e. y because $T_x M$, $d^M f_x$ exist for $\mathcal{H}^n$-a.e. $x \in M_+$), we have

3.7 $$\frac{d^M f_{x\#}\xi(x)}{|d^M f_{x\#}\xi(x)|} = \pm\tau_1 \wedge \cdots \wedge \tau_n,$$

where $\tau_1, \ldots, \tau_n$ is an orthonormal basis for $T_y(f(M))$. Hence (a) gives

$$f_\# T(\omega) = \int_{f(M)} \langle \omega(y), \eta(y)\rangle N(y)\, d\mathcal{H}^n(y)$$

where $\eta(y)$ is a suitable orientation for the approximate tangent space $T_y(f(M))$ and $N(y)$ is a non-negative integer. N, η in fact satisfy

3.8 $$\textstyle\sum_{x \in f^{-1}(y)\cap M_+} \theta(x) \dfrac{d^M f_{x\#}\xi(x)}{|d^M f_{x\#}\xi(x)|} = N(y)\,\eta(y),$$

so that for $\mathcal{H}^n$-a.e. $y \in f(M)$ we have

$$N(y) \le \textstyle\sum_{x\in f^{-1}(y)\cap M_+} \theta(x)$$

and

$$N(y) \equiv \textstyle\sum_{x\in f^{-1}(y)\cap M_+} \theta(x) \pmod 2.$$

Thus we have proved

3.9 Lemma. *If $f : U \to W$ is locally Lipschitz and $f|\operatorname{spt} T$ is proper, with $T = \tau(M, \xi, \theta) \in \mathcal{D}_n(U)$ an integer multiplicity current, then $f_\# T$ is an integer multiplicity current in W; in fact $f_\# T = \tau(f(M), \eta, N)$, where η, N are as in 3.8 above.*

Notice in particular by applying 3.9 to the current $R = 0 ⨳ T$ in 2.32, we have

3.10 U star-shaped relative to 0, $T \in \mathcal{D}_n(U)$ integer multiplicity, $\operatorname{spt} T$ compact, and $\partial T = 0 \Rightarrow \exists$ an integer multiplicity $R \in \mathcal{D}_{n+1}(U)$ with $\partial R = T$.

Observe that, in case f is C^1, $f_\# T$ agrees with the previous definition in 2.23 (see also 2.24(2)). Notice also that if $f : U \to W$ is Lipschitz and if $V = \underline{v}(M, \theta)$ is the varifold associated with $T = \underline{\tau}(M, \xi, \theta)$, then

$$\mu_{f_\# T} \le \mu_{f_\# V}$$

(in the sense of measures) with equality if and only if, for $\mathcal{H}^n$-a.e. $y \in f(M)$, the sign in 3.7 above remains constant as x varies over $f^{-1}(y) \cap M_+$. In particular we have $\mu_{f_\# T} = \mu_{f_\# V}$ in case f is 1:1.

A fact of central importance concerning integer multiplicity currents is the following compactness theorem, first proved by Federer and Fleming [FF60]:

3.11 Theorem. (Federer-Fleming Compactness Theorem.) *If $\{T_j\} \subset \mathcal{D}_n(U)$ is a sequence of integer multiplicity currents with*

$$\sup_{j\ge1}\big(\mathbb{M}_W(T_j) + \mathbb{M}_W(\partial T_j)\big) < \infty \ \forall W \subset\subset U,$$

then there is an integer multiplicity $T \in \mathcal{D}_n(U)$ and a subsequence $\{T_{j'}\}$ such that $T_{j'} \rightharpoonup T$ in U.

We shall give the proof of this in §8. Notice that the existence of a $T \in \mathcal{D}_n(U)$ and a subsequence $\{T_{j'}\}$ with $T_{j'} \rightharpoonup T$ is a consequence of the elementary 2.15; only the fact that T is an integer multiplicity current is non-trivial.

3.12 Remark: Notice that the proof of 3.11 in the codimension 1 case (when $P = n$) is a direct consequence of Remark 2.35 and the Compactness Theorem for BV functions (§2.6 of Ch. 2).

In contrast to the difficulty in proving 3.11, it is quite straightforward to prove that if T_j converges to T in the strong sense that $\lim \mathbb{M}_W(T_j - T) = 0\ \forall W \subset\subset U$, and if T_j are integer multiplicity $\forall j$, then T is integer multiplicity. Indeed we have the following lemma.

3.13 Lemma. *The set of integer multiplicity currents in $\mathcal{D}_n(U)$ is complete with respect to the topology given by the family $\{\mathbb{M}_W\}_{W\subset\subset U}$ of semi-norms.*

Proof: Let $\{T_Q\}$ be a sequence of integer multiplicity currents in $\mathcal{D}_n(U)$ and $\{T_Q\}$ is Cauchy with respect to the semi-norms $\mathbb{M}_W$, $W \subset\subset U$. Suppose

$$T_Q = \underline{\tau}(M_Q, \theta_Q, \xi_Q)$$

(θ_Q positive integer-valued on M_Q, M_Q countably n-rectifiable, $\mathcal{H}^n(M_Q \cap W) < \infty$ for each $W \subset\subset U$). Then

(1) $$\mathbb{M}_W(T_Q - T_P) \equiv \int_W |\theta_P \xi_P - \theta_Q \xi_Q|\, d\mathcal{H}^n < \varepsilon_W(Q)$$

$\forall P \ge Q$, where $\varepsilon_W(Q) \downarrow 0$ as $Q \to \infty$ and where we adopt the convention $\xi_P = 0$, $\theta_P = 0$ on $U \setminus M_P$. In particular, since $|\xi_P| = 1$ on M_P, we get

(2) $$\int_W |\theta_P - \theta_Q|\, d\mathcal{H}^n < \varepsilon_W(Q)\ \forall P \ge Q,$$

and hence θ_P converges in $L^1(\mathcal{H}^n)$ locally in U to an integer-valued function θ. Of course (2) implies

(3) $$\mathcal{H}^n(((M_+ \setminus M_Q) \cup (M_Q \setminus M_+)) \cap W) \le \varepsilon_W(Q),$$

where $M_+ = \{x \in U : \theta(x) > 0\}$. (1), (2) also imply

$$\int_W \theta_P |\xi_P - \xi_Q| \, d\mathcal{H}^n \le 2\varepsilon_W(Q) \ \forall P \ge Q,$$

and hence by (3) ξ_P converges in $L^1(\mathcal{H}^n)$ locally in U to a function ξ with values in $\Lambda_n(\mathbb{R}^{n+\ell})$ with $|\xi| = 1$ and ξ simple on M_+.

Now $\xi_q(x) \in \Lambda_n(T_x M_Q)$, $\mathcal{H}^n$-a.e. $x \in M_Q$, and (by (3)) $T_x M_+ = T_x M_Q$ except for a set of measure $\le \varepsilon_W(Q)$ in $M_+ \cap W$. It follows that $\xi(x) \in \Lambda_n(T_x M_+)$ for $\mathcal{H}^n$-a.e. $x \in M_+$ and we have shown that $\mathbb{M}_W(T_P - T) \to 0$, where $T = \underline{\tau}(M_+, \theta, \xi)$ is an integer multiplicity n-current in U. □

Finally, we shall need the following useful *decomposition theorem* for codimension 1 integer multiplicity currents.

3.14 Theorem. *Suppose $P = n+1$ (i.e. U is open in $\mathbb{R}^{n+1}$) and R is an integer multiplicity current in $\mathcal{D}_{n+1}(U)$ with $\mathbb{M}_W(\partial R) < \infty \ \forall W \subset\subset U$. Then $T = \partial R$ is integer multiplicity, and in fact we can find a decreasing sequence of $\mathcal{L}^{n+1}$-measurable sets $\{U_j\}_{j=-\infty}^{\infty}$ of locally finite perimeter in U such that (in U)*

$$\begin{aligned} R &= \textstyle\sum_{j=1}^{\infty} [\![U_j]\!] - \sum_{j=-\infty}^{0} [\![V_j]\!], \quad V_j = U \setminus U_j, \quad j \le 0, \\ T &= \textstyle\sum_{j=-\infty}^{\infty} \partial [\![U_j]\!], \\ \mu_T &= \textstyle\sum_{j=-\infty}^{\infty} \mu_{\partial [\![U_j]\!]}, \end{aligned}$$

and in particular

$$\mathbb{M}_W(T) = \textstyle\sum_{j=-\infty}^{\infty} \mathbb{M}_W(\partial [\![U_j]\!]) \ \forall W \subset\subset U.$$

3.15 Remark: Let $* : C_c^\infty(U; \mathbb{R}^{n+1}) \to \mathcal{D}^n(U)$ be defined by

$$*g = \textstyle\sum_{j=1}^{n+1} (-1)^{j-1} g_j \, dx^1 \wedge \cdots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \cdots \wedge dx^{n+1},$$

so that $d * g = \operatorname{div} g \, dx^1 \wedge \cdots \wedge dx^{n+1}$. Then for any $\mathcal{L}^{n+1}$-measurable $A \subset U$ we have

$$\begin{aligned} \partial [\![A]\!] (*g) &= [\![A]\!] (d * g) \\ &= \int_U \chi_A \operatorname{div} g \, d^{n+1}, \end{aligned}$$

and hence by definition of $|D\chi_A|$ (in §2 of Ch. 2) and $\mathbb{M}(T)$ (in §2 of the present chapter) we see that

$$A \text{ has locally finite perimeter in } U \iff \mathbb{M}_W(\partial [\![A]\!]) < \infty \ \forall W \subset\subset U,$$

and in this case

$$\begin{cases} \mathbb{M}_W(\partial [\![A]\!]) = \int_W |D\chi_A| \ \forall W \subset\subset U \\ \overrightarrow{\partial [\![A]\!]} = *\nu_A, \ |D\chi_A| \text{ a.e. in } U. \end{cases}$$

Here ν_A is the inward unit normal function for A (defined on the reduced boundary $\partial^* A$ as in 4.3 of Ch. 3).

Proof of 3.14: R must have the form

$$R = \underline{\tau}(V, \theta, \xi),$$

where V is an $\mathcal{L}^{n+1}$-measurable subset of U and $\xi(x) = \pm e_1 \wedge \cdots \wedge e_{n+1}$ for each $x \in V$. Thus letting

$$\widetilde{\theta}(x) = \begin{cases} \theta(x) & \text{when } x \in V \text{ and } \xi(x) = +e_1 \wedge \cdots \wedge e_{n+1} \\ -\theta(x) & \text{when } x \in V \text{ and } \xi(x) = -e_1 \wedge \cdots \wedge e_{n+1} \\ 0 & \text{when } x \notin V, \end{cases}$$

we have

(1) $$R(\omega) = \int_V a \widetilde{\theta} \, d\mathcal{L}^{n+1},$$

$\omega = a \, dx^1 \wedge \cdots \wedge dx^{n+1} \in \mathcal{D}^{n+1}(U)$ and (cf. 2.8)

(2) $$\mathbb{M}_W(R) = \int_W |\widetilde{\theta}| \, d\mathcal{L}^{n+1}, \ \mathbb{M}_W(T) = \int_W |D\widetilde{\theta}| \ \forall W \subset\subset U$$

(and $\widetilde{\theta} \in BV_{\text{loc}}(U)$).

Define

$$\begin{aligned} U_j &= \left\{x \in U : \widetilde{\theta}(x) \ge j\right\}, \ j \in \mathbb{Z} \\ V_j &= U \setminus U_j = \left\{x \in U : \widetilde{\theta}(x) \le -1 - j\right\}, \ j \le 0. \end{aligned}$$

Then one checks directly that

$$\widetilde{\theta} = \textstyle\sum_{j=1}^{\infty} \chi_{U_j} - \sum_{j=-\infty}^{0} \chi_{V_j}$$

(χ_A = indicator function of A, $A \subset U$), and hence by (1)

(3) $$R = \textstyle\sum_{j=1}^{\infty} [\![U_j]\!] - \sum_{j=-\infty}^{0} [\![V_j]\!] \text{ in } U.$$

Since $T(\omega) = \partial R(\omega) = R(d\omega)$, $\omega \in \mathcal{D}^n(U)$, we then have

(4) $$\begin{aligned} T = \partial R &= \textstyle\sum_{j=1}^{\infty} \partial [\![U_j]\!] - \sum_{j=-\infty}^{0} \partial [\![V_j]\!] \\ &= \textstyle\sum_{j=-\infty}^{\infty} \partial [\![U_j]\!], \end{aligned}$$

so we have the required decomposition, and it remains only to prove that each U_j has locally finite perimeter in U and that the corresponding measures add. To check this, take $\psi_j \in C^1(\mathbb{R})$ with

$$\begin{cases} \psi_j(t) = 0 \quad \text{for} \quad t \le j-1+\varepsilon, \quad \psi_j(t) = 1, \quad t \ge j-\varepsilon \\ 0 \le \psi_j \le 1, \quad \sup|\psi_j'| \le 1+3\varepsilon, \end{cases}$$

where $\varepsilon \in (0, \frac{1}{2})$. Then if $a \in C_c^\infty(U)$ and $g = (g^1, \ldots, g^{n+1})$, $g^j \in C_c^\infty(U)$, with $|g| \le a$, we have (since $\chi_{U_j} = \psi_j \circ \widetilde{\theta}\ \forall j$) that for any $M \le N$

$$\begin{aligned} (5) \qquad \int_U \operatorname{div} g \textstyle\sum_{j=M}^N \chi_{U_j}\, d\mathcal{L}^{n+1} &= \int_U \operatorname{div} g \textstyle\sum_{j=M}^N \psi_j \circ \widetilde{\theta}\, d\mathcal{L}^{n+1} \\ &= \lim_{\sigma\downarrow 0} \int_U \operatorname{div} g \textstyle\sum_{j=M}^N \psi_j \circ \widetilde{\theta}_\sigma\, d\mathcal{L}^{n+1} \\ &= -\lim_{\sigma\downarrow 0} \int_U g \cdot \nabla\widetilde{\theta}_\sigma\, \psi_j'(\widetilde{\theta}_\sigma)\, d\mathcal{L}^{n+1} \\ &\le (1+3\varepsilon) \lim_{\sigma\downarrow 0} \int_U a|\nabla\widetilde{\theta}_\sigma|\, d\mathcal{L}^{n+1}. \end{aligned}$$

On the other hand

$$(6) \qquad \begin{cases} \int_U a|\nabla\widetilde{\theta_\sigma}| = \sup_{g\in C_c^1(U), |g|\le a} \int_U \operatorname{div} g\, \widetilde{\theta}_\sigma\, d\mathcal{L}^n, \text{ and} \\ \int_U \operatorname{div} g\, \widetilde{\theta}_\sigma\, d\mathcal{L}^n = \int_U \operatorname{div} g_\sigma\, \widetilde{\theta}\, d\mathcal{L}^n = R(d\omega_\sigma) = T(\omega_\sigma) \le \mathbb{M}(T)|\omega|, \end{cases}$$

where $\omega_\sigma = \sum_{j=1}^n (-1)^{j-1} g_{j\sigma} dx^1 \wedge \cdots \wedge dx^{j-1} \wedge dx^{j+1} \wedge \cdots \wedge dx^n$. Thus, taking $M = N$, we deduce from (5) and (6) that $\mathbb{M}_W(\partial[\![U_j]\!]) \le \mathbb{M}_W(T) < \infty$ for each j and each open $W \subset\subset U$.

By taking $M = -N$ in (5) and defining $R_N = \sum_{j=1}^N [\![U_j]\!] - \sum_{j=-N}^0 [\![V_j]\!]$ we see that (with g as in 3.15)

$$\left|R_N(d * g)\right| \le (1+3\varepsilon) \int_U a\, d\mu_T,$$

and hence, with $T_N = \partial R_N$,

$$(7) \qquad \int_U a\, d\mu_{T_N} \le \int_U a\, d\mu_T \ \ \forall N \ge 1,$$

$a \ge 0$, $a \in C_c^\infty(U)$. On the other hand by 4.1 of Ch. 3 we have

$$\begin{aligned} R_N(d * g) &= \textstyle\sum_{j=-N}^N \int_U \operatorname{div}\, g\, \chi_{U_j}\, d\mathcal{L}^{n+1} \\ (8) \qquad &= -\textstyle\sum_{j=-N}^N \int_{\partial^* U_j} \nu_j \cdot g\, d\mathcal{H}^n, \end{aligned}$$

where ν_j is the inward unit normal for U_j and $\partial^* U_j$ is the reduced boundary for U_j (see §4 of Ch. 3 and in particular 4.3 of Ch. 3). By virtue of the fact that $U_{j+1} \subset U_j$ we see from 4.3 (‡‡) of Ch. 3 that $\nu_j \equiv \nu_k$ on $\partial^* U_j \cap \partial^* U_k\ \forall j, k$. Hence (8) can be written

$$T_N(*g) = -\int_U \nu \cdot g\, h_N\, d\mathcal{H}^n,$$

where $h_N = \sum_{j=-N}^N \chi_{\partial^* U_j}$ and where ν is defined on $\cup_{j=-\infty}^\infty \partial^* U_j$ by $\nu = \nu_j$ on $\partial^* U_j$. Since $|\nu| \equiv 1$ on $\cup_{j=-\infty}^\infty \partial^* U_j$ this evidently gives

$$\begin{aligned} \int a\, d\mu_{T_N} &= \int a\, h_N\, d\mathcal{H}^n \\ &= \textstyle\sum_{j=-N}^N \int_{\partial^* U_j} a\, d\mathcal{H}^n \\ &= \textstyle\sum_{j=-N}^N \int a\, d\mu_{\partial[\![U_j]\!]}. \end{aligned}$$

Letting $N \to \infty$ we thus have (by (7))

$$\mu_T \ge \textstyle\sum_{j=-\infty}^\infty \mu_{\partial[\![U_j]\!]}.$$

Since the reverse inequality follows directly from (4), the proof is complete. □

3.16 Corollary. *Let R be integer multiplicity $\in \mathcal{D}_{n+1}(U)$, $U \subset \mathbb{R}^P$, $P \ge n+1$, and suppose there is an $(n+1)$-dimensional C^1 submanifold N of $\mathbb{R}^P$ with* spt $R \subset N \cap U$. *Suppose further that $T = \partial R$ and $\mathbb{M}(T) < \infty\ \forall W \subset\subset U$. Then T ($\in \mathcal{D}_n(U)$) is integer multiplicity and for each point $y \in N \cap U$ there is an open $W_y \subset\subset U$, $y \in W_y$, and $\mathcal{H}^{n+1}$ measurable subset $\{U_j\}_{j=-\infty}^\infty$ with $U_{j+1} \subset U_j \subset N \cap U$, $\mathbb{M}_{W_y}(\partial[\![U_j]\!]) < \infty$ $\forall j$, and with the following identities holding in W_y:*

$$\begin{aligned} R &= \textstyle\sum_{j=1}^\infty [\![U_j]\!] - \sum_{j=-\infty}^0 [\![U \setminus U_j]\!] \\ T &= \textstyle\sum_{j=-\infty}^\infty \partial[\![U_j]\!] \\ \mu_T &= \textstyle\sum_{j=-\infty}^\infty \mu_{\partial[\![U_j]\!]}. \end{aligned}$$

Proof: The proof is an easy consequence of 3.14 using local coordinate representations for N. □

4 Slicing

We first want to define the notion of slice for integer multiplicity currents. Preparatory to this we have the following lemma:

4.1 Lemma. *If M is $\mathcal{H}^n$-measurable, countably n-rectifiable, f is Lipschitz on $\mathbb{R}^{n+\ell}$ and $M_+ = \{x \in M : |\nabla^M f(x)| > 0\}$, then for $\mathcal{L}^1$-almost all $t \in \mathbb{R}$ the following statements hold:*

(1) *$M_t \equiv f^{-1}(t) \cap M_+$ is countably $\mathcal{H}^{n+1}$-rectifiable*

(2) *For $\mathcal{H}^{n-1}$-a.e. $x \in M_t$, $T_x M_t$ and $T_x M$ both exist, $T_x M_t$ is an $(n-1)$-dimensional subspace of $T_x M$, and in fact*

$$(\ddagger) \qquad T_x M = \{y + \lambda \nabla^M f(x) : y \in T_x M_t, \lambda \in \mathbb{R}\}.$$

Furthermore for any non-negative $\mathcal{H}^n$-measurable function g on M we have

$$\int_{-\infty}^{\infty}\Big(\int_{M_t} g\, d\mathcal{H}^{n-1}\Big)\, dt = \int_M |\nabla^M f|\, g\, d\mathcal{H}^n.$$

Proof: In fact (1) is just a restatement of 2.10(2) of Ch. 3, and (2) follows from 1.6 of Ch. 3 together with the facts that for $\mathcal{L}^1$-a.e. $t \in \mathbb{R}$ and $\mathcal{H}^{n-1}$-a.e. $x \in M_t$

$$\nabla^M f(x) \in T_x M \quad \text{(by definition of } \nabla^M f \text{ in §2 of Ch. 3)}$$

and

$$\langle \nabla^M f(x), \tau\rangle = 0 \;\forall \tau \in T_x M_t.$$

(This last follows for example from Definition 2.1 of Ch. 3.)

The last part of the lemma is just a restatement of the appropriate version of the co-area formula (discussed in §2 of Ch. 3).

4.2 Remark: Note that by replacing g (in 4.1 above) by $g\times$characteristic function of $\{x : f(x) < t\}$ we get the identity

$$\int_{M\cap\{f(x)<t\}} |\nabla^M f| g\, d\mathcal{H}^n = \int_{-\infty}^{t}\int_{M_s} g\, d\mathcal{H}^{n-1} ds$$

so that the left side as an absolutely continuous function of t and

$$\frac{d}{dt}\int_{M\cap\{f(x)<t\}} |\nabla^M f| g\, d\mathcal{H}^n = \int_{M_t} g\, d\mathcal{H}^{n-1}, \text{ a.e.} t \in \mathbb{R}.$$

Now let $T = \underline{\tau}(M, \theta, \xi)$ be an integer multiplicity current in U (U open in $\mathbb{R}^{n+\ell}$, $M \subset U$), let f be Lipschitz in U and let θ_+ be defined $\mathcal{H}^n$-a.e. in M by

$$\theta_+(x) = \begin{cases} 0 & \text{if } \nabla^M f(x) = 0 \\ \theta(x) & \text{if } \nabla^M f(x) \neq 0. \end{cases}$$

For the ($\mathcal{L}^1$-almost all) $t \in \mathbb{R}$ such that $T_x M$, $T_x M_t$ exist for $\mathcal{H}^{n-1}$-a.e. $x \in M_t$ and such that 4.1 (2)($\ddagger$) holds, we have

$$\textbf{4.3} \qquad \xi(x) \llcorner \frac{\nabla^M f(x)}{|\nabla^M f(x)|} \textit{ is simple } \in \Lambda_{n-1}(T_x M_t) \subset \Lambda_{n-1}(T_x M)$$

and has unit length (for $\mathcal{H}^{n-1}$-a.e. $x \in M_t$). Here we use the notation that if $v \in \Lambda_n(T_x M)$ and $w \in T_x M$, then $v \llcorner w \in \Lambda_{n-1}(T_x M)$ is defined by

$$\langle v \llcorner w, a\rangle = \langle v, w \wedge a\rangle,\ a \in \Lambda_{n-1}(T_x M).$$

Using this notation we can now define the notion of a slice of T by f; we continue to assume $T \in \mathcal{D}_n(U)$ is given by $T = \underline{\tau}(M, \theta, \xi)$ as above.

4.4 Definition: For the ($\mathcal{L}^1$-almost all) $t \in \mathbb{R}$ since that $T_x M$, $T_x M_t$ exist and Lemma 4.1 (2)($\ddagger$) holds $\mathcal{H}^{n-1}$-a.e. $x \in M_t$, with the notation introduced above (and bearing in mind 4.3) we define the integer multiplicity current $\langle T, f, t\rangle \in \mathcal{D}_{n-1}(U)$ by

$$\langle T, f, t\rangle = \underline{\tau}(M_t, \theta_t, \xi_t),$$

where

$$\xi_t(x) = \xi(x) \llcorner \frac{\nabla^M f(x)}{|\nabla^M f(x)|},\ \theta_t = \theta_+|M_t.$$

So defined, $\langle T, f, t\rangle$ is called the slice of T by f at t.

4.5 Lemma. (1) *For each open $W \subset U$*

$$\int_{-\infty}^{\infty} \mathbb{M}_W(\langle T, f, t\rangle)\, dt = \int_{M\cap W} |\nabla^M f| \theta\, d\mathcal{H}^n \le \big(\operatorname*{ess\,sup}_{M\cap W} |\nabla^M f|\big)\mathbb{M}_W(T).$$

(2) *If $\mathbb{M}_W(\partial T) < \infty\ \forall W \subset\subset U$, then for $\mathcal{L}^1$-a.e. $t \in \mathbb{R}$*

$$\langle T, f, t\rangle = \partial[T \llcorner \{f < t\}] - (\partial T) \llcorner \{f < t\}.$$

(3) *If ∂T is integer multiplicity in $\mathcal{D}_{n-1}(U)$, then for $\mathcal{L}^1$-a.e. $t \in \mathbb{R}$*

$$\langle \partial T, f, t\rangle = -\partial\langle T, f, t\rangle.$$

Proof: (1) is a direct consequence of the last part of 4.1 (with $g = \theta_+$).

To prove (2) we first recall that, since M is countably n-rectifiable, we can write (see Remark 1.3 of Ch. 3)

$$M = \cup_{j=0}^{\infty} M_j,$$

where $M_i \cap M_j = \emptyset\ \forall i \neq j$, $\mathcal{H}^n(M_0) = 0$, and $M_j \subset N_j$ $j \ge 1$, with N_j an embedded C^1 submanifold of $\mathbb{R}^{n+\ell}$. By virtue of this decomposition and the definition of ∇^M (in §2 of Ch. 3) it easily follows that if h is Lipschitz on $\mathbb{R}^{n+\ell}$ and if $h^{(\sigma)}$ are the mollified functions (as in §2 of Ch. 2) then, as $\sigma \downarrow 0$,

(1) $$v \cdot \nabla^M h^{(\sigma)} \to v \cdot \nabla^M h \quad \text{(weak convergence in } L^2(\mu_T))$$

for any fixed bounded $\mathcal{H}^n$-measurable v with values in $\mathbb{R}^{n+\ell}$. (Indeed to check this, we have merely to check that (1) holds with N_j in place of M_j and with v vanishing on $\mathbb{R}^{n+\ell} \setminus M_j$; since N_j is C^1 this follows fairly easily by approximating v by smooth functions and using the fact that $h^{(\sigma)}$ converges to h uniformly.)

Next let $\varepsilon > 0$ and let γ be the Lipschitz function on $\mathbb{R}$ defined by

$$\gamma(s) = \begin{cases} 1, & s < t - \varepsilon \\ \text{linear}, & t - \varepsilon \le s \le t \\ 0, & s > t \end{cases}$$

and apply the above to $h = \gamma \circ f$. Then letting $\omega \in \mathcal{D}^n(U)$ we have

$$\begin{aligned} \partial T\left(h^{(\sigma)}\omega\right) &= T\left(d\left(h^{(\sigma)}\omega\right)\right) \\ &= T\left(dh^{(\sigma)} \wedge \omega\right) + T\left(h^{(\sigma)} d\omega\right). \end{aligned}$$

Then using the integral representations of the form 2.7 for ∂T we see that

(2) $$(\partial T \llcorner h)(\omega) = \lim_{\sigma \downarrow 0} T\left(dh^{(\sigma)} \wedge \omega\right) + (T \llcorner h)(d\omega).$$

Since $\xi(x)$ orients $T_x M$, we have

$$\begin{aligned} \left\langle \xi(x), dh^{(\sigma)} \wedge \omega \right\rangle &= \left\langle \xi(x), \left(dh^{(\sigma)}(x)\right)^T \wedge \omega^T \right\rangle \\ &= \left\langle \xi(x), \left(dh^{(\sigma)}(x)\right)^T \wedge \omega \right\rangle \end{aligned}$$

(where $(\ )^T$ denotes the orthogonal projection of $\Lambda^q(\mathbb{R}^{n+\ell})$ onto $\Lambda^q(T_x M)$). Thus

$$\begin{aligned} T\left(dh^{(\sigma)} \wedge \omega\right) &= \int_M \left\langle \xi(x), \left(dh^{(\sigma)}(x)\right)^T \wedge \omega \right\rangle \theta \, d\mathcal{H}^n \\ &= \int_M \left\langle \xi(x) \llcorner \nabla^M h^{(\sigma)}(x), \omega \right\rangle \theta \, d\mathcal{H}^n \end{aligned}$$

so that by (1)

(3) $$\lim_{\sigma \downarrow 0} T\left(dh^{(\sigma)} \wedge \omega\right) = \int_M \left\langle \xi(x) \llcorner \nabla^M h(x), \omega \right\rangle \theta \, d\mathcal{H}^n.$$

By 2.1 of Ch. 3 of $\nabla^M h$ and by the chain rule for the composition of Lipschitz functions we have

(4) $$\nabla^M h = \gamma'(f) \nabla^M f \quad \mathcal{H}^n\text{-a.e. on } M$$

(where we set $\gamma'(f) = 0$ when f takes the "bad" values t or $t - \varepsilon$; note that $\nabla^M h(x) = \nabla^M f(x) = 0$ for $\mathcal{H}^n$-a.e. in $\{x \in M : f(x) = c\}$, c any given constant).

Using (3), (4) in (2), we thus deduce

$$(\partial T \llcorner h)(\omega) = -\varepsilon^{-1} \int_{M \, \{t - \varepsilon < f < t\}} \left\langle \xi \llcorner \nabla^M f, \omega \right\rangle \theta \, d\mathcal{H}^n + (T \llcorner h)(d\omega).$$

Finally we let $\varepsilon \downarrow 0$ and we use 4.2 with $g = \theta \left\langle \xi \llcorner \frac{\nabla^M f}{|\nabla^M f|}, \omega \right\rangle$ in order to complete the proof of (2); by considering a countable dense set of $\omega \in \mathcal{D}^n(U)$ one can of course show that 4.2 is applicable with $g = \theta \left\langle \xi \llcorner \frac{\nabla^M f}{|\nabla^M f|}, \omega \right\rangle$ except for a set F of t having $\mathcal{L}^1$-measure zero, with F independent of ω.

Finally to prove part (3) of the theorem, we first apply part (2) with ∂T in place of T. Since $\partial^2 T = 0$, this gives

$$\langle \partial T, f, t \rangle = \partial \left[(\partial T) \llcorner \{f < t\} \right].$$

On the other hand, applying ∂ to each side of the original identity (for T) of (2), we get

$$\partial \left[(\partial T) \llcorner \{f < t\} \right] = -\partial \langle T, f, t \rangle$$

and hence (3) is established. □

Motivated by the above discussions we are led to define slices for an arbitrary current $\in \mathcal{D}_n(U)$ which, together with its boundary, has locally finite mass in U. Specifically, suppose $\mathbb{M}_W(T) + \mathbb{M}_W(\partial T) < \infty \ \forall W \subset\subset U$. Then we define "slices"

4.6 $$\langle T, f, t_- \rangle = \partial(T \llcorner \{f < t\}) - (\partial T) \llcorner \{f < t\}$$

and

4.7 $$\langle T, f, t_+ \rangle = -\partial(T \llcorner \{f > t\}) + (\partial T) \llcorner \{f > t\}.$$

Of course $\langle T, f, t_+ \rangle = \langle T, f, t_- \rangle$ (and the common value is denoted $\langle T, f, t \rangle$) for all but the countably many values of t such that $\mathbb{M}(T \llcorner \{f = t\}) + \mathbb{M}((\partial T) \llcorner \{f = t\}) > 0$.

The important properties of the above slices are that if f is Lipschitz on U (and if we continue to assume $\mathbb{M}_W(T) + \mathbb{M}_W(\partial T) < \infty \ \forall W \subset\subset U$), then

4.8 $$\operatorname{spt} \langle T, f, t_+ \rangle \subset \operatorname{spt} T \cap \{x : f(x) = t\}$$

and, $\forall$ open $W \subset U$,

4.9
$$\begin{cases} \mathbb{M}_W(\langle T, f, t_+\rangle) \le \operatorname{ess\,sup}_W |Df| \liminf_{h\downarrow 0} h^{-1}\mathbb{M}_W(T \llcorner \{t < f < t+h\}) \\ \mathbb{M}_W(\langle T, f, t_-\rangle) \le \operatorname{ess\,sup}_W |Df| \liminf_{h\downarrow 0} h^{-1}\mathbb{M}_W(T \llcorner \{t-h < f < t\}). \end{cases}$$

Notice that $\mathbb{M}_W(T \llcorner \{f < t\})$ is increasing in t, hence is differentiable for $\mathcal{L}^1$-a.e. $t \in \mathbb{R}$ and $\int_a^b \frac{d}{dt}\mathbb{M}_W(T \llcorner \{f < t\})\, dt \le \mathbb{M}_W(T \llcorner \{a < f < b\})$. Thus 4.9 gives

4.10
$$\int_a^{*b} \mathbb{M}_W(\langle T, f, t_\pm\rangle)\, dt \le \operatorname{ess\,sup}_W |Df|\ \mathbb{M}(T \llcorner \{a < f < b\})$$

for every open $W \subset U$.

To prove 4.8 and 4.9 we consider first the case when f is C^1 and take any smooth increasing function $\gamma : \mathbb{R} \to \mathbb{R}_+$ and note that

4.11
$$\begin{aligned} \partial(T \llcorner \gamma\circ f)(\omega) - ((\partial T)\llcorner \gamma\circ f)(\omega) &= (T \llcorner \gamma \circ f)(d\omega) - ((\partial T)\llcorner \gamma\circ f)(\omega) \\ &= T(\gamma\circ f\, d\omega) - T(d(\gamma\circ f\omega)) \\ &= -T(\gamma'(f)\, df \wedge \omega). \end{aligned}$$

Now let $\varepsilon > 0$ be arbitrary and choose γ such that

$$\gamma(t) = 0 \text{ for } t < a,\ \gamma(t) = 1 \text{ for } t > b,\ 0 \le \gamma'(t) \le \frac{1+\varepsilon}{b-a} \text{ for } a < t < b.$$

Then the left side of 4.11 converges to $\langle T, f, a_+\rangle$ if we let $b \downarrow a$, and hence 4.8 follows because $\operatorname{spt}\gamma' \subset [a,b]$. Furthermore the right side R of 4.11 evidently satisfies

$$|R| \le (\sup_W |Df|)\Big(\frac{1+\varepsilon}{b-a}\Big)\, \mathbb{M}_W(T \llcorner \{a < f < b\})|\omega| \qquad (\operatorname{spt}\omega \subset W)$$

and so we also conclude the first part of 4.11 for $f \in C^1$. We similarly establish the second part for $f \in C^1$. To handle general Lipschitz f we simply use $f^{(\sigma)}$ in place of f in 4.6, 4.7 and in the above proof, then let $\sigma \downarrow 0$ where appropriate.

5 The Deformation Theorem

The deformation theorem, given below in 5.1 and 5.3, is a central result in the theory of currents, and was first proved by Federer and Fleming [FF60].

The special notation for this section is as follows:

$n, k \in \{1, 2, \ldots\}$,

$C = [0,1] \times \cdots \times [0,1]$ (standard unit cube in $\mathbb{R}^{n+\ell}$)

$\mathbb{Z}^{n+\ell} = \{z = (z^1, \ldots, z^{n+\ell}) : z^j \in \mathbb{Z}\}$ (integer lattice in $\mathbb{R}^{n+\ell}$)

L_j = j-skeleton of the decomposition $\cup_{z\in\mathbb{Z}^{n+\ell}}(z + C)$ of $\mathbb{R}^{n+\ell}$

$\mathcal{L}_j$ = collection of j-faces in L_j
$= \{z + F : z \in \mathbb{Z}^{n+\ell},\ F$ is a closed j-face of $C\}$

$\mathcal{L}_j(\rho) = \{\rho F : F \in \mathcal{L}_j\}$, $\rho > 0$

$S_1, \ldots, S_N$ $\left(N = \binom{n+\ell}{n+1} = \binom{n+\ell}{\ell-1}\right)$ denote the $(n+1)$-dimensional subspaces of $\mathbb{R}^{n+\ell}$ which contain an $(n+1)$-face of the standard cube C.

p_j denotes the orthogonal projection of $\mathbb{R}^{n+\ell}$ onto S_j, $j = 1, \ldots, n$.

5.1 (Deformation Theorem, unscaled version.) *Suppose T is an n-current in $\mathbb{R}^{n+\ell}$ (i.e. $T \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$) with $\mathbb{M}(T) + \mathbb{M}(\partial T) < \infty$. Then we can write*

$$T - P = \partial R + S$$

where P, R, S satisfy

$$P = \textstyle\sum_{F\in\mathcal{L}_n} \beta_F [\![F]\!] \quad (\beta_F \in \mathbb{R}),$$

with

$$\mathbb{M}(P) \le C\mathbb{M}(T),\ \mathbb{M}(\partial P) \le C\mathbb{M}(\partial T)$$
$$\mathbb{M}(R) \le C\mathbb{M}(T),\ \mathbb{M}(S) \le C\mathbb{M}(\partial T)$$

($C = C(n,k)$), and

$$\operatorname{spt} P \cup \operatorname{spt} R \subset \left\{x : \operatorname{dist}(x, \operatorname{spt} T) < 2\sqrt{n+\ell}\right\}$$
$$\operatorname{spt}\partial P \cup \operatorname{spt} S \subset \left\{x : \operatorname{dist}(x, \operatorname{spt}\partial T) < 2\sqrt{n+\ell}\right\}.$$

In case T is an integer multiplicity current, then P, R can be chosen to be integer multiplicity currents (and the β_F appearing in the definition of P are integers). If in addition ∂T is integer multiplicity[2], then S can be chosen to be integer multiplicity.

5.2 Remarks: (1) Note that this is slightly sharper than the corresponding theorem in [FF60], [Fed69], because there is no term involving $\mathbb{M}(\partial T)$ in the bound for $\mathbb{M}(P)$.

(2) It follows automatically from the other conclusions of the theorem that $\mathbb{M}(\partial S) \le C\mathbb{M}(\partial T)$. Also, it follows from the inequalities $\mathbb{M}(\partial P)$, $\mathbb{M}(S) \le C\mathbb{M}(\partial T)$ that $S = 0$ and $\partial P = 0$ when $\partial T = 0$.

[2] Actually if $\mathbb{M}(\partial T) < \infty$ then ∂T is *automatically* integer multiplicity if T is—see 6.3 below.

The following "scaled version" of 5.1 is obtained from the above by first changing scale $s \to \rho^{-1}x$, then applying 5.1, then changing scale back by $x \to \rho x$.

5.3 (Deformation Theorem, scaled version.) *Suppose T, ∂T are as in* 5.1, *and $\rho > 0$. Then*

$$T - P = \partial R + S$$

where P, R, S satisfy

$$P = \textstyle\sum_{F \in \mathcal{L}_j(\rho)} \beta_F [\![F]\!] \qquad (\beta_F \in \mathbb{R})$$

$$\mathbb{M}(P) \le C\mathbb{M}(T),\ \mathbb{M}(\partial P) \le C\mathbb{M}(\partial T)$$

$$\mathbb{M}(R) \le C\rho\mathbb{M}(T),\ \mathbb{M}(S) \le C\rho\mathbb{M}(\partial T),$$

and

$$\operatorname{spt} P \cup \operatorname{spt} R \subset \left\{x : \operatorname{dist}(x, \operatorname{spt} T) < 2\sqrt{n+\ell}\,\rho\right\}$$

$$\operatorname{spt} \partial P \cup \operatorname{spt} S \subset \left\{x : \operatorname{dist}(x, \operatorname{spt} \partial T) < 2\sqrt{n+\ell}\,\rho\right\}.$$

As in 5.1, *in case T is integer multiplicity, so are P, R; if ∂T is integer multiplicity then so is s.*

The main step in the proof of the deformation theorem will involve "pushing" T onto the n-skeleton L_n via a certain retraction map ψ. We first have to establish the existence of a suitable class of retraction maps. This is done in the following lemma, in which we use the notation

$$q = \text{center point of } C = \left(\tfrac{1}{2}, \tfrac{1}{2}, \dots, \tfrac{1}{2}\right),$$

$$L_{k-1}(a) = a + L_{k-1}\ (a \text{ a given point in } B_{1/4}(q)),$$

$$L_{k-1}(a;\rho) = \left\{x \in \mathbb{R}^{n+\ell} : \operatorname{dist}\big(x, L_{k-1}(a)\big) < \rho\right\}\ \left(\rho \in \left(0, \tfrac{1}{4}\right)\right).$$

Note that $\operatorname{dist}\big(L_{k-1}(a), L_n\big) \ge \frac{1}{4}$ for any point $a \in B_{1/4}(q)$.

5.4 Lemma. *For every $a \in B_{\frac{1}{4}}(q)$ there is a locally Lipschitz map*

$$\psi : \mathbb{R}^{n+\ell} \setminus L_{k-1}(a) \to \mathbb{R}^{n+\ell} \setminus L_{k-1}(a)$$

such that

$$\psi(C \setminus L_{k-1}(a)) = C \cap L_n,\ \psi | C \cap L_n = 1_{C \cap L_n},$$

$$|D\psi(x)| \le \frac{c}{\rho},\ \mathcal{L}^{n+\ell}\text{-a.e. } x \in C \setminus L_{k-1}(a;\rho),\ \rho \in \left(0, \frac{1}{4}\right),$$

$(c = c(n,k))$, and such that

$$\psi(z + x) = z + \psi(x),\ x \in \mathbb{R}^{n+\ell} \setminus L_{k-1}(a),\ z \in \mathbb{Z}^{n+\ell}.$$

Proof: We first construct a locally Lipschitz retraction $\psi_0 : C \setminus L_{k-1}(a)$ onto the n-faces of C. This is done as follows:

Firstly for each j-face F of C, $j \ge n+1$, let $a_F \in F$ denote the orthogonal projection of a onto F, and let ψ_F denote the retraction of $\overline{F} \setminus \{a_F\}$ onto ∂F which takes a point $x \in \overline{F} \setminus \{a_F\}$ to the point $y \in \partial F$ such that $x \in \{a_F + \lambda(y - a_F) : \lambda \in (0,1]\}$. (Thus ψ_F is the "radial retraction" of F with a_F as origin.) Of course $\psi_F | \partial F = 1_{\partial F}$. Notice also that for any j-face F of C, $j \ge n+1$, the line segment $\overline{aa}_F$ is contained in $L_{k-1}(a)$; in fact if J_F denotes the set of ℓ such that S_ℓ (see notation prior to 5.1) is parallel to an $(n+1)$-face of F, then (because $\overline{aa}_F$ is orthogonal to F, hence orthogonal to each S_ℓ, $\ell \in J_F$) we have

$$\text{(1)} \qquad \overline{aa}_F \subset \cap_{\ell \in J_F} p_\ell^{-1}(p_\ell(a)),$$

and this is contained in $L_{k-1}(a)$, because (by definition)

$$\text{(2)} \qquad L_{k-1}(a) = \cup_{\ell=1}^{N} \cup_{z \in \mathbb{Z}^{n+\ell}} \left(z + p_\ell^{-1}(p_\ell(a))\right).$$

Next, for each $j \ge n+1$, define

$$\psi^{(j)} : \cup \left\{\overline{F} \setminus \{a_F\} : F \text{ is a } j\text{-face of } C\right\} \to \cup \left\{\overline{G} : G \text{ is a } (j-1)\text{-face of } C\right\}$$

by setting

$$\psi^{(j)} \big| \overline{F} \setminus \{a_F\} = \psi_F.$$

(Notice that then $\psi^{(j)}$ is locally Lipschitz on its domain by virtue of the fact that each ψ_F is the identity on ∂F, F any j-face of C.)

Then the composite $\psi^{(n+1)} \circ \psi^{(n+2)} \circ \dots \circ \psi^{(n+\ell)}$ makes sense on $C \setminus L_{k-1}(a)$ (by (1)), so we can set

$$\psi_0 = \psi^{(n+1)} \circ \psi^{(n+2)} \circ \dots \circ \psi^{(n+\ell)} \big| C \setminus L_{k-1}(a).$$

Notice that ψ_0 has the additional property that if

$$z \in \mathbb{Z}^{n+\ell} \text{ and } x,\ z + x \in C, \text{ then } \psi_0(z + x) = z + \psi_0(x).$$

(Indeed $x, z + x \in C$ means that either $x, z + x$ are in L_n (where ψ_0 is the identity) or else lie in the interior of parallel j-faces F_1, $F_2 = z + F_1$ $(j \ge n+1)$ of C with z orthogonal to F_1 and $a_{F_2} = z + a_{F_1}$.) It follows that we can then define a retraction ψ of all of $C \setminus L_{k-1}(a)$ onto L_n by setting

$$\psi(z + x) = z + \psi_0(x),\ x \in C \setminus L_{k-1}(a),\ z \in \mathbb{Z}^{n+\ell}.$$

We now claim that

(3) $$\sup|D\psi| \le \frac{c}{\rho} \text{ on } \mathbb{R}^{n+\ell} \setminus L_{k-1}(a,\rho),\ c = c(n,k).$$

(This will evidently complete the proof of the lemma.)

We can prove (3) by induction on k as follows. First note that (3) is evident from construction in case $\ell = 1$. Hence assume $k \ge 2$ and assume (3) holds in case $\ell - 1$ replaces ℓ in the above construction. Let x be any point of interior $(C)\setminus L_{\ell-1}(a;\rho)$, let $y = \psi^{n+\ell}(x)$ ($\psi^{n+\ell}$ is the radial retraction of $C \setminus \{a\}$ onto ∂C, and let F be any closed $(n+\ell-1)$-face of C which contains y.

Suppose now new coordinates are selected so that $F \subset \mathbb{R}^{n+\ell-1} \times \{0\} \subset \mathbb{R}^{n+\ell}$, and also let $\widetilde{L}_{k-2}(a) = L_{k-1}(a) \cap \mathbb{R}^{n+\ell-1} \times \{0\}$). By virtue of (1) we have $a_F \in L_{k-1}(a)$, hence

(4) $$|y - a_F| \ge \operatorname{dist}(y, L_{k-1}(a)).$$

Let p_F be the orthogonal projection of $\mathbb{R}^{n+\ell}$ onto $\mathbb{R}^{n+\ell-1} \times \{0\}$ ($\supset F$), so that $a_F = p_F(a)$. Evidently $|p_F(x) - a_F| \ge \operatorname{dist}(x, p_F^{-1}(p_F(a)))$ and hence by (2) we deduce

(5) $$|p_F(x) - a_F| \ge \operatorname{dist}(x, L_{k-1}(a)).$$

Furthermore by definition of y we know that $y - a = \frac{|y-a|}{|x-a|}(x-a)$ and hence, applying p_F, we have

$$y - a_F = \frac{|y-a|}{|x-a|}p_F(x-a).$$

Hence since $|y-a| \ge 3/4$, we have

(6) $$|y - a_F| \ge \tfrac{3}{4}\frac{|p_F(x-a)|}{|x-a|}.$$

Now let $\widetilde{\psi}$ be the retraction of $F \setminus \widetilde{L}_{k-2}(a)$ onto the n-faces of F ($\widetilde{\psi}$ defined as for ψ but with $(k-1)$ in place of k, a_F in place of a, $\mathbb{R}^{n+\ell-1}$ in place of $\mathbb{R}^{n+\ell}$ and $\widetilde{L}_{k-2}(a) = L_{k-2}(a_F)$ in place of $L_{k-1}(a)$). By the inductive hypothesis, together with (4), (5), (6) we have

$$\begin{aligned}\left|\overline{D}\widetilde{\psi}(y)\right| &\le \frac{c}{\operatorname{dist}(y, \widetilde{L}_{k-2}(a))},\ \Big(\left|\overline{D}\widetilde{\psi}(y)\right| = \limsup_{z\to y}\frac{\left|\widetilde{\psi}(z) - \widetilde{\psi}(y)\right|}{|z-y|}\Big) \\ &\le \left(\tfrac{4}{3}\right)c\frac{1}{|y-a_F|}\frac{|x-a|}{|p_F(x-a)|} \\ &\le \left(\tfrac{4}{3}\right)c\frac{|x-a|}{\operatorname{dist}(x, L_{k-1}(a))},\end{aligned} \tag{7}$$

when $k = 2$. For general k, we label $L = L_{k-2}(a_F)$ and note that $\frac{\operatorname{dist}(y,L)}{\operatorname{dist}\left(x, p_F^{-1}(L)\right)} = \frac{|y-a|}{|x-a|}$ by similarity, and $p_F^{-1}(L) \subset L_{k-1}(a)$. So $|\overline{D}\widetilde{\psi}(y)| \le \frac{c|x-a|}{\operatorname{dist}(x, L_{k-1}(a))}$ as required. Also, by the definition of $\psi^{n+\ell}$ we have that

(8) $$\left|\overline{D}\psi^{n+\ell}(x)\right| \le \frac{c}{|x-a|},\ \left|\overline{D}\psi^{n+\ell}(x)\right| = \limsup_{y\to x}\frac{\left|\psi^{n+\ell}(y) - \psi^{n+\ell}(x)\right|}{|y-x|}.$$

Since $\psi(x) = \widetilde{\psi}\circ\psi^{n+\ell}(x)$, we have by (7), (8) and the chain rule that

$$\begin{aligned}\left|\overline{D}\psi(x)\right| \le \left|\overline{D}\widetilde{\psi}(y)\right|\left|\overline{D}\psi^{n+\ell}(x)\right| &\le \frac{c}{|x-a|}\frac{|x-a|}{\operatorname{dist}(x, L_{k-1}(a))} \\ &= \frac{c}{\operatorname{dist}(x, L_{k-1}(a))}. \quad \square\end{aligned}$$

Proof of the Deformation Theorem:

We use the subspaces $S_1, \ldots, S_N$ and projections $p_1, \ldots, p_N$ introduced at the beginning of the section. Let $F_j = C \cap S_j$ (so that F_j is a closed $(n+1)$-dimensional face of C), let x_j by the central point of F_j, and for each $j = 1, \ldots, N$ define a "good" subset $G_j \subset F_j \cap B_{\frac{1}{4}}(x_j)$ by $g \in G_j \iff g \in F_j \cap B_{\frac{1}{4}}(x_j)$ and

(1) $$\mathbb{M}\big(T \llcorner \cup_{z\in\mathbb{Z}^{n+\ell}\cap S_j}\, p_j^{-1}\big(B_\rho(g+z)\big)\big) \le \beta\rho^{n+1}\mathbb{M}(T) \qquad \forall\rho \in \big(0, \frac{1}{4}\big)$$

(β to be chosen, $G_j = G_j(\beta)$).

We now claim that the "bad" set $B_j = F_j \cap B_{\frac{1}{4}}(x_j) \setminus G_j$ in fact has $\mathcal{L}^{n+1}$-measure (taken in S_j) small; in fact we claim

(2) $$\mathcal{L}^{n+1}(B_j) \le 20^{n+1}\beta^{-1}\omega_{n+1}\big(\tfrac{1}{4}\big)^{n+1} \quad (\omega_{n+1} = \mathcal{L}^{n+1}(B_1(0))),$$

which is indeed small if we choose large β. To see (2), we argue as follows. For each $b \in B_j$ there is (by definition) a $\rho_b \in (0, \frac{1}{4})$ such that

(3) $$\mathbb{M}\big(T \llcorner \cup_{z\in\mathbb{Z}^{n+\ell}\cap S_j}\, p_j^{-1}\big(B_{\rho_b}(b+z)\big)\big) > \beta\rho_b^{n+1}\mathbb{M}(T),$$

and by the 5-times Covering Lemma 3.4 of Ch. 1 there is a pairwise disjoint subcollection $\{B_{\rho_\ell}(b_\ell)\}_{\ell=1,2,\ldots}$ ($\rho_\ell = \rho_{b_\ell}$) of the collection $\{B_{\rho_b}(b)\}_{b\in B_j}$ such that

(4) $$B_j \subset \cup_\ell B_{5\rho_\ell}(b_\ell).$$

But then, setting $b = b_\ell$ in (3) and summing, we get

$$\beta\big(\textstyle\sum_\ell\rho_\ell^{n+1}\big)\,\mathbb{M}(T) \le \mathbb{M}(T) \quad (\text{i.e. } \textstyle\sum_\ell\rho_\ell^{n+1} \le \beta^{-1}),$$

(using the fact that $\left\{p_j^{-1}B_{\rho_\ell}(b_\ell+z)\right\}_{\ell=1,2,\dots z\in\mathbb{Z}^{n+\ell}\cap S_j}$ is a pairwise disjoint collection for fixed j). Thus by (4) we conclude

$$\mathcal{L}^{n+1}(B_j)\le \beta^{-1}5^{n+1}\omega_{n+1},$$

which after trivial re-arrangement gives (2) as required. Thus we have

$$\mathcal{L}^{n+1}(G_j)\ge \left(1-20^{n+1}\beta^{-1}\right)\omega_{n+1}\left(\tfrac14\right)^{n+1},$$

and it follows that

$$(5)\qquad \mathcal{L}^{n+\ell}\left(p_j^{-1}(G_j)\cap B_{\frac14}(q)\right)\ge \left(1-\frac{\omega_{n+1}}{\omega_{n+\ell}}20^{n+1}\beta^{-1}\right)\omega_{n+\ell}\left(\tfrac14\right)^{n+\ell},$$

where q is the center point $\left(\frac12,\dots,\frac12\right)$ of C. (So $p_j(q)=x_j$.)

Then selecting β large enough so that $20^{n+1}\omega_{n+1}N\beta^{-1}<\omega_{n+\ell}/(n+\ell)$, we see from (5) that we can choose a point $a\in\cap_{j=1}^N p_j^{-1}(G_j)\cap B_{\frac14}(q)$. Next let $L_{k-1}(a)=a+L_{k-1}$, $L_{k-1}(a;\rho)=\left\{x\in\mathbb{R}^{n+\ell}:\operatorname{dist}(x,L_{k-1}(a))<\rho\right\}$ (as in the proof of 5.4) and note that in fact

$$L_{k-1}(a;\rho)=\cup_{j=1}^N\ \cup_{z\in\mathbb{Z}^{n+\ell}\cap S_j}\ p_j^{-1}(B_\rho(p_j(a)+z)).$$

Then since $p_j(a)\in G_j$ we have (by definition of G_j)

$$(6)\qquad \mathbb{M}\left(T\llcorner L_{k-1}(a;\rho)\right)\le N\beta\rho^{n+1}\mathbb{M}(T)\qquad \forall\rho\in\left(0,\tfrac14\right).$$

Indeed let us suppose that we take β such that $20^{n+1}\omega_{n+1}N\beta^{-1}<\omega_{n+\ell}/2(n+\ell)$. Then more than half the ball $B_{\frac14}(q)$ is in the set $\cap_{j=1}^N p_j^{-1}(G_j)$ and hence, repeating the whole argument above with ∂T in place of T, we can actually select a so that, *in addition to* (6), we also have

$$(7)\qquad \mathbb{M}\left(\partial T\llcorner L_{k-1}(a;\rho)\right)\le N\beta\rho^{n+1}\mathbb{M}(\partial T)\qquad \forall\rho\in\left(0,\tfrac14\right).$$

Now let ψ be the retraction of $\mathbb{R}^{n+\ell}\setminus L_{k-1}(a)$ onto L_n given in 5.4, and let

$$(8)\qquad T_\rho=T\llcorner L_{k-1}(a;\rho),\ (\partial T)_\rho=\partial T\llcorner L_{k-1}(a;\rho),$$

so that by (6), (7)

$$(9)\qquad \mathbb{M}(T_\rho)\le c\rho^{n+1}\mathbb{M}(T),\ \mathbb{M}((\partial T)_\rho)\le c\rho^{n+1}\mathbb{M}(\partial T),\ \forall\rho\in\left(0,\tfrac14\right).$$

Furthermore by 4.10 we know that for each $\rho\in\left(0,\frac14\right)$ we can find $\rho^*\in(\rho/2,\rho)$ such that

$$(10)\qquad \mathbb{M}\left(\langle T,d,\rho^*\rangle\right)\le\frac{c}{\rho}\mathbb{M}(T_\rho-T_{\rho/2})\le c\rho^n\mathbb{M}(T),$$

where d is the (Lipschitz) distance function to $L_{k-1}(a)$ ($d(x)=\operatorname{dist}(x,L_{k-1}(a))$, $\operatorname{Lip}(d)=1$) and $\langle T,d,\rho^*\rangle$ is the slice of T by d at ρ^*. (Notice that we can choose ρ^* such that (10) holds and such that $\langle T,d,\rho^*\rangle$ is integer multiplicity—see 4.5 and the following discussion.)

We now want to apply the homotopy formula 2.25 to the case when $h(x,t)=x+t(\psi(x)-x)$, $\in\mathbb{R}^{n+\ell}\setminus L_{k-1}(a;\sigma)$, $\sigma>0$. Notice that h is Lipschitz on $\mathbb{R}^{n+\ell}\setminus L_{k-1}(a;\sigma)$, so we can define $h_\#$, $\psi_\#$ as in 2.29. (We shall apply $h_\#$, $\psi_\#$ only to currents supported away from $[0,1]\times L_{k-1}(a)$ and $L_{k-1}(a)$ respectively.)

Keeping this in mind we note that by 5.4, (6) and (7) we have

$$(11)\qquad \mathbb{M}(\psi_\#(T_\rho-T_{\rho/2}))\le\frac{c}{\rho^n}\rho^{n+1}\mathbb{M}(T)\le c\rho\mathbb{M}(T)$$

and

$$(12)\qquad \mathbb{M}(\psi_\#((\partial T)_\rho-(\partial T)_{\rho/2}))\le\frac{c}{\rho^{n-1}}\rho^{n+1}\mathbb{M}(\partial T)\le c\rho\mathbb{M}(\partial T).$$

Similarly by the homotopy formula 2.25, together with 2.27 and (6), (7) above, we have

$$(13)\qquad \mathbb{M}(h_\#([\![(0,1)]\!]\times(T_\rho-T_{\rho/2})))\le c\rho\mathbb{M}(T)$$

and

$$(14)\qquad \mathbb{M}(h_\#([\![(0,1)]\!]\times((\partial T)_\rho-(\partial T)_{\rho/2}))\le c\rho\mathbb{M}(\partial T).$$

Notice also that by (6), (10) and 2.27 we have

$$(15)\qquad \mathbb{M}\left(\psi_\#\langle T,d,\rho^*\rangle\right)\le c\rho\mathbb{M}(T)$$

and

$$(16)\qquad \mathbb{M}\left(h_\#\left([\![(0,1)]\!]\times\langle T,d,\rho^*\rangle\right)\right)\le c\rho\mathbb{M}(T).$$

Next note that by iteration (11), (12) imply

$$(17)\qquad \begin{cases}\mathbb{M}(\psi_\#(T_\rho-T_{\rho/2^\nu}))\le 2c\rho\mathbb{M}(T)\\ \mathbb{M}(\psi_\#((\partial T)_\rho-(\partial T)_{\rho/2^\nu}))\le 2c\rho\mathbb{M}(\partial T)\end{cases}$$

for each integer $\nu\ge1$, where c is as in (11), (12) (c independent of ν). Selecting $\rho=\frac14$ and using the arbitrariness of ν, it follows that

$$(18)\qquad \begin{cases}\mathbb{M}(\psi_\#(T-T_\sigma))\le c\mathbb{M}(T)\\ \mathbb{M}(\psi_\#(\partial T-(\partial T)_\sigma))\le c\mathbb{M}(\partial T)\end{cases}$$

for each $\sigma \in (0,1)$ (with c independent of σ).

Now select $\rho = \rho_\nu = 2^{-\nu}$ and $\rho_\nu^* \in [2^{-\nu-1}, 2^{-\nu}]$ such that (10), (15), (16) hold with ρ_ν^* in place of ρ^*; then by (15), (16), (17), (18) we have that

$$\psi_\#(T - T_{\rho_\nu^*}),\ h_\#([\![(0,1)]\!] \times (T - T_{\rho_\nu^*})),$$
$$\psi_\#(\partial T - \partial T_{\rho_\nu^*}),\ h_\#([\![(0,1)]\!] \times \partial(T - T_{\rho_\nu^*}))$$

are Cauchy sequences relative to $\mathbb{M}$, and $\mathbb{M}(\langle T, d, \rho_\nu^*\rangle) + \mathbb{M}(\psi_\# \langle T, d, \rho_\nu^*\rangle) \to 0$. Hence there are currents $Q, S_1 \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$ and $R_1 \in \mathcal{D}_{n+1}(\mathbb{R}^{n+\ell})$ such that

$$\text{(19)} \qquad \begin{cases} \lim \mathbb{M}(Q - \psi_\#(T - T_{\rho_\nu^*})) = 0 \\ \lim \mathbb{M}(S_1 - h_\#([\![(0,1)]\!] \times \partial(T - T_{\rho_\nu^*}))) = 0 \\ \lim \mathbb{M}(R_1 - h_\#([\![(0,1)]\!] \times (T - T_{\rho_\nu^*}))) = 0. \end{cases}$$

Furthermore by the homotopy formula and 2.27 we have for each ν

$$\text{(20)} \qquad \begin{aligned} T - T_{\rho_\nu^*} - \psi_\#(T - T_{\rho_\nu^*}) &= \partial(h_\#([\![(0,1)]\!] \times (T - T_{\rho_\nu^*}))) \\ &\quad - h_\#([\![(0,1)]\!] \times \partial(T - T_{\rho_\nu^*})). \end{aligned}$$

Since $\partial T_{\rho_\nu^*} = (\partial T)_{\rho_\nu^*} - \langle T, d, \rho_\nu^*\rangle$ (by the definition 4.6, 4.7 of slice) we thus get that

$$\text{(21)} \qquad T - Q = \partial R_1 + S_1.$$

(Notice that Q, R_1 are integer multiplicity by (19), 4.4, 4.5 and 3.13 in case T is integer multiplicity; similarly S_1 is integer multiplicity if ∂T is.)

Using the fact that ψ retracts $\mathbb{R}^{n+\ell} \setminus L_{k-1}(a)$ onto L_n we know (by 2.27) that $\operatorname{spt} \psi_\#(T - T_{\rho_\nu^*}) \subset L_n$, and hence

$$\text{(22)} \qquad \operatorname{spt} Q \subset L_n.$$

We also have (since $\psi(z + C) \subset z + C\ \forall z \in \mathbb{Z}^{n+\ell}$) that

$$\text{(23)} \qquad \begin{cases} \operatorname{spt} R_1 \cup \operatorname{spt} Q \subset \{x : \operatorname{dist}(x, \operatorname{spt} T) < \sqrt{n+\ell}\} \\ \operatorname{spt} S_1 \subset \{x : \operatorname{dist}(x, \operatorname{spt} \partial T) < \sqrt{n+\ell}\} \end{cases}$$

and, by (18), (19), we have

$$\text{(24)} \qquad \mathbb{M}(Q) \le c\mathbb{M}(T),\ \mathbb{M}(R_1) \le c\mathbb{M}(T),\ \mathbb{M}(S_1) \le c\mathbb{M}(\partial T).$$

Also by (18) and the semi-continuity of $\mathbb{M}$ under weak convergence, we have

$$\text{(25)} \qquad \begin{aligned} \mathbb{M}(\partial Q) &\le \liminf \mathbb{M}(\partial \psi_\#(T - T_{\rho_\nu^*})) \\ &= \liminf \mathbb{M}(\psi_\# \partial(T - T_{\rho_\nu^*})) \\ &\le c\mathbb{M}(\partial T). \end{aligned}$$

Now let F be a given face of L_n(i.e. $F \in \mathcal{L}_n$) and let $\breve{F}$ = interior of F. Assume for the moment that $F \subset \mathbb{R}^n \times \{0\}$ ($\subset \mathbb{R}^{n+\ell}$), and let p be the orthogonal projection onto $\mathbb{R}^n \times \{0\}$. By construction of ψ we know that $p \circ \psi = \psi$ in a neighborhood of any point $y \in \breve{F}$. We therefore have (since Q is given by (18)) that

$$\text{(26)} \qquad p_\#(Q \llcorner \breve{F}) = Q \llcorner \breve{F}.$$

It then follows, by the obvious modifications of the arguments in the proof of the Constancy Theorem (2.34) and in 2.35, that

$$\text{(27)} \qquad (Q \llcorner \breve{F}) = \int_{\breve{F}} \langle e_1 \wedge \cdots \wedge e_n, \omega(x)\rangle \theta_F(x)\, d\mathcal{L}^n(x)$$

$\forall \omega \in \mathcal{D}^n(\mathbb{R}^{n+\ell})$, for some $BV_{\text{loc}}(\mathbb{R}^n)$ function θ_F, and

$$\text{(28)} \qquad \mathbb{M}(Q \llcorner \breve{F}) = \int_{\breve{F}} |\theta_F|\, d\mathcal{L}^n,\ \mathbb{M}((\partial Q) \llcorner \breve{F}) = \int_{\breve{F}} |D\theta_F|.$$

Furthermore, since

$$\text{(29)} \qquad (Q \llcorner \breve{F} - \beta[\![F]\!])(\omega) = \int_{\breve{F}} (\theta_F - \beta)\, \langle e_1 \wedge \cdots \wedge e_n, \omega(x)\rangle\, d\mathcal{L}^n(x)$$

(by (27)), we have (again using the reasoning of 2.35)

$$\text{(30)} \qquad \begin{cases} \mathbb{M}(Q \llcorner \breve{F} - \beta[\![F]\!]) = \int_{\breve{F}} |\theta_F - \beta|\, d\mathcal{L}^n \\ \mathbb{M}(\partial(Q \llcorner \breve{F} - \beta[\![F]\!])) = \int_{\mathbb{R}^n} |D(\chi_{\breve{F}}(\theta_F - \beta))|, \end{cases}$$

where $\chi_{\breve{F}}$ = characteristic function of $\breve{F}$. Thus taking $\beta = \beta_F$ such that

$$\text{(31)} \qquad \min\left\{\mathcal{L}^n\left\{x \in \breve{F} : \theta_F \ge \beta\right\},\ \mathcal{L}^n\left\{x \in \breve{F} : \theta_F(x) \le \beta\right\}\right\} \ge \frac{1}{2}$$

(which we can do because $\mathcal{L}^n(\breve{F}) = 1$; notice that we can, and we do, take $\beta_F \in \mathbb{Z}$ if θ_F is integer-valued), we have by 2.7 and 2.9 of Ch. 2, (28) and (30) that

$$\text{(32)} \qquad \begin{cases} \mathbb{M}(Q \llcorner \breve{F} - \beta[\![F]\!]) \le c \int_{\breve{F}} |D\theta_F| = c\mathbb{M}(\partial Q \llcorner \breve{F}) \\ \mathbb{M}(\partial(Q \llcorner \breve{F} - \beta[\![F]\!])) \le c \int_{\breve{F}} |D\theta_F| = c\mathbb{M}(\partial Q \llcorner \breve{F}). \end{cases}$$

We also have by 2.37(1)

$$\text{(33)} \qquad Q \llcorner \partial F = 0.$$

Then summing over $F \in \mathcal{L}_n$ and using (32), (33) we have, with $P = \sum_{F\in\mathcal{L}_n}\beta_F[\![F]\!]$, that

$$(34)\qquad \begin{cases} \mathbb{M}(Q-P) \le c\mathbb{M}(\partial Q) \\ \mathbb{M}(\partial Q - \partial P) \le c\mathbb{M}(\partial Q). \end{cases}$$

Actually by (31) we have

$$(35)\qquad |\beta_F| \le 2\int_{\check F}|\theta_F|\,d\mathcal{L}^n,$$

and hence (using again the first part of (28)), since $\mathbb{M} = \sum_F|\beta_F|$,

$$(36)\qquad \mathbb{M}(P) \le c\mathbb{M}(Q).$$

Notice that the second part of (34) gives

$$(37)\qquad \mathbb{M}(\partial P) \le c\mathbb{M}(\partial Q).$$

Finally we note that (21) can be written

$$(38)\qquad T - P = \partial R_1 + \big(S_1 + (Q-P)\big).$$

Setting $R = R_1$, $S = S_1 + (Q-P)$, the theorem now follows immediately from (23), (24), (25) and (34), (36), (37), (38); the fact that P, R are integer multiplicity if T is should be evident from the remarks during the course of the above proof, as should be the fact that S is integer multiplicity if $T, \partial T$ are. □

6 Applications of the Deformation Theorem

We here establish a couple of simple (but very important) applications of the deformation theorem, namely the isoperimetric theorem and the weak polyhedral approximation theorem. This latter theorem, when combined with the compactness 3.11, implies the important "Boundary Rectifiability Theorem" (6.3 below), which asserts that if T is an integer multiplicity current in $\mathcal{D}_n(U)$ and if $\mathbb{M}_W(\partial T) < \infty$ $\forall W \subset\subset U$, then $\partial T (\in \mathcal{D}_{n-1}(U))$ is integer multiplicity. (Notice that in the case $k = 0$, this has already been established in 3.14.)

6.1 (Isoperimetric Theorem.) *Suppose $T \in \mathcal{D}_{n-1}(\mathbb{R}^{n+\ell})$ is integer multiplicity, $n \ge 2$, spt T is compact and $\partial T = 0$. Then there is an integer multiplicity current $R \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$ with spt R compact, $\partial R = T$, and*

$$\mathbb{M}(R)^{\frac{n-1}{n}} \le c\mathbb{M}(T),$$

where $c = c(n,k)$.

Proof: The case $T = 0$ is trivial, so assume $T \ne 0$. Let P, R, S be integer multiplicity currents as in 5.3, where for the moment $\rho > 0$ is arbitrary, and note that $S = 0$ because $\partial T = 0$. Evidently (since $\mathcal{H}^{n-1}(F) = \rho^{n-1}$ $\forall F \in \mathcal{F}_{n-1}(\rho)$) we have

$$(1)\qquad \mathbb{M}(P) = N(\rho)\rho^{n-1}$$

for some non-negative *integer* $N(\rho)$. But since $\mathbb{M} \le c\mathbb{M}(T)$ (from 5.3) we see that necessarily $N(\rho) = 0$ in (1) if we choose $\rho = \big(2c\mathbb{M}(T)\big)^{\frac{1}{n-1}}$. Then $P = 0$, and 5.3 gives $T = \partial R$ for some integer multiplicity current R with spt R compact and $\mathbb{M}(R) \le c\rho\mathbb{M}(T) = c'\big(\mathbb{M}(T)\big)^{\frac{1}{n-1}}$. □

6.2 (Weak Polyhedral Approximation Theorem.) *Given any integer multiplicity $T \in \mathcal{D}_n(U)$ with $\mathbb{M}_W(T), \mathbb{M}_W(\partial T) < \infty$ $\forall W \subset\subset U$, there is a sequence $\{P_k\}$ of current of the form*

$$(\ddagger)\qquad P_k = \textstyle\sum_{F\in\mathcal{F}_n(\rho_k)}\beta_R^{(k)}[\![F]\!], \quad (\beta_R^{(k)} \in \mathbb{Z}),\ \rho_k \downarrow 0,$$

such that $P_k \rightharpoonup T$ (and hence also $\partial P_k \rightharpoonup \partial T$) in U (in the sense of 2.13).

Proof: First consider the case $U = \mathbb{R}^{n+\ell}$ and $\mathbb{M}(T)$, $\mathbb{M}(\partial T) < \infty$. In this case we simply use the deformation theorem: for any sequence $\rho_k \downarrow 0$, the scaled version of the deformation theorem (with $\rho = \rho_k$) gives P_k as in (‡) such that

$$(1)\qquad T - P_k = \partial R_k + S_k$$

for some R_k, S_k such that

$$(2)\qquad \begin{cases} \mathbb{M}(R_k) \le c\rho_k\mathbb{M}(T) \to 0 \\ \mathbb{M}(S_k) \le c\rho_k\mathbb{M}(\partial T) \to 0 \end{cases}$$

and

$$\mathbb{M}(P_k) \le c\mathbb{M}(T),\ \mathbb{M}(\partial P_k) \le c\mathbb{M}(\partial T).$$

Evidently (1), (2) give $P_k(\omega) \to T_k(\omega)$ $\forall \omega \in \mathcal{D}^n(\mathbb{R}^{n+\ell})$, and $\partial P_k = 0$ if $\partial T = 0$, so the theorem is proved in case $U = \mathbb{R}^{n+\ell}$ and $T, \partial T$ are of finite mass.

In the general case we take any Lipschitz function φ on $\mathbb{R}^{n+\ell}$ such that $\varphi > 0$ in U, $\varphi \equiv 0$ in $\mathbb{R}^{n+\ell} \setminus U$ ans such that $\{x = \varphi(x) > \lambda\} \subset\subset U$ $\forall \lambda > 0$. For $\mathcal{L}^1$-a.e. $\lambda > 0$, 4.5 implies that $T_\lambda \equiv T \llcorner \{x : \varphi(x) > \lambda\}$ is such that $\mathbb{M}(\partial T_\lambda) < \infty$. Since spt $T_\lambda \subset\subset U$, we can apply the argument above to approximate T_λ for any such λ.

Taking a suitable sequence $\lambda_j \downarrow 0$, the required approximation then immediately follows. □

6.3 (Boundary Rectifiability Theorem.) *Suppose T is an integer multiplicity current in $\mathcal{D}_n(U)$ with $\mathbb{M}(\partial T) < \infty\ \forall W \subset\subset U$. Then $\partial T (\in \mathcal{D}_{n-1}(U))$ is an integer multiplicity current.*

Proof: A direct consequence of 6.2 above and the Compactness 3.11.

6.4 Remark: Notice that only the case $\partial T_j = 0\ \forall j$ of 3.11 is needed in the above proof.

7 The Flat Metric Topology

The main result to be proved here is the equivalence of weak convergence and "flat metric"[3] convergence (see below for terminology) for a sequence of integer multiplicity currents $\{T_j\} \subset \mathcal{D}_n(U)$ such that $\sup_{j\ge 1}(\mathbb{M}_W(T_j)+\mathbb{M}_W(\partial T_j)) < \infty\ \forall W \subset\subset U$.

We let U denote (as usual) an arbitrary open subset of $\mathbb{R}^{n+\ell}$,

$$\mathcal{I} = \{T \in \mathcal{D}_n(U) : T \text{ is integer multiplicity and } \mathbb{M}_W(\partial T) < \infty\ \forall W \subset\subset U\}.$$

and

$$\mathcal{I}_{M,W} = \{T \in \mathcal{I} : \operatorname{spt} T \subset \overline{W},\ \mathbb{M}(T) + \mathbb{M}(\partial T) \le M\}.$$

for any $M > 0$ and open $W \subset\subset U$.

On $\mathcal{I}$ we define a family of pseudometrics $\{d_W\}_{W\subset\subset U}$ by

7.1 $$d_W(T_1,T_2) = \inf\{\mathbb{M}_W(S) + \mathbb{M}_W(R) : T_1 - T_2 = \partial R + S, \text{ where } R \in \mathcal{D}_{n+1}(U),\ S \in \mathcal{D}_n(U) \text{ are integer multiplicity}\}.$$

We henceforth assume $\mathcal{I}$ is equipped with the topology given (in the usual way) by the family $\{d_W\}_{W\subset\subset U}$ of pseudometrics. This topology is called the "flat metric topology" for $\mathcal{I}$: there is a countable base of neighborhoods at each point, and $T_j \to T$ in this topology if and only if $d_W(T_j,T) \to 0\ \forall W \subset\subset U$.

7.2 Theorem. *Let T, $\{T_j\} \subset \mathcal{D}_n(U)$ be integer multiplicity with*

$$\sup_{j\ge 1}(\mathbb{M}_W(T_j) + \mathbb{M}_W(\partial T_j)) < \infty\ \forall W \subset\subset U.$$

[3] Note that the word "flat" here has *no* physical or geometric significance, but relates rather to Whitney's use of the symbol ♭ (the "flat" symbol in musical notation) in his work. We mention this because it is often a source of confusion.

Then $T_j \rightharpoonup T$ (in the sense of 2.13) if and only if $d_W(T_j,T) \to 0$ for each $W \subset\subset U$.

7.3 Remark: Notice that no use is made of the Compactness 3.11 in this theorem; however if we combine the compactness theorem with it, then we get the statement that for any family of positive (finite) constants $\{c(W)\}_{W\subset\subset U}$ the set $\{T \in \mathcal{I} : \mathbb{M}_W(T_j) + \mathbb{M}_W(\partial T_j) \le c(W)\ \forall W \subset\subset U\}$ is sequentially compact when equipped with the flat metric topology.

Proof of 7.2: First note that the "if" part of the theorem is trivial (indeed for a given $W \subset\subset U$, the statement $d_W(T_j,T) \to 0$ evidently implies $(T_j - T)(\omega) \to 0$ for any fixed $\omega \in \mathcal{D}^n(U)$ with $\operatorname{spt}\omega \subset W$).

For the "only if" part of the theorem, the main difficulty is to establish the appropriate "total boundedness" property; specifically we show that for any given $\varepsilon > 0$ and $W \subset\subset \widetilde{W} \subset\subset U$, we can find $N = N(\varepsilon, W, \widetilde{W}, M)$ and integer multiplicity currents $P_1,\ldots,P_N \in \mathcal{D}_n(U)$ such that

$$\text{(1)} \qquad \mathcal{I}_{M,W} \subset \textstyle\sum_{j=1}^N B_{\varepsilon,\widetilde{W}}(P_j),$$

where, for any $P \in \mathcal{I}$, $B_{\varepsilon,\widetilde{W}}(P) = \{S \in \mathcal{I} : d_{\widetilde{W}}(S,P) < \varepsilon\}$. This is an easy consequence of the Deformation Theorem: in fact for any $\rho > 0$, 5.3 guarantees that for $T \in \mathcal{I}_{M,W}$ we can find integer multiplicity P, R, S such that

$$\text{(2)} \qquad \begin{cases} T - P = \partial R + S \\ P = \sum_{F\in\mathcal{F}_n(\rho)} \beta_F [\![F]\!],\ \beta_F \in \mathbb{Z} \\ \operatorname{spt} P \subset \{x : \operatorname{dist}(x, \operatorname{spt} T) < 2\sqrt{n+\ell}\,\rho\} \end{cases}$$

$$\text{(3)} \qquad \begin{cases} \mathbb{M}(P)\ (\equiv \sum_{F\in\mathcal{F}_n(\rho)} |\beta_F|\rho^n) \le c\mathbb{M}(T) \le cM \\ \operatorname{spt} R \cup \operatorname{spt} S \subset \{x : \operatorname{dist}(x, \operatorname{spt} T) < 2\sqrt{n+\ell}\,\rho\} \\ \mathbb{M}(R) + \mathbb{M}(S) \le c\rho\mathbb{M}(T) \le c\rho M. \end{cases}$$

Then for ρ small enough to ensure $2\sqrt{n+\ell}\rho < \operatorname{dist}(W, \partial\widetilde{W})$, we see from (2),(3) that

$$d_{\widetilde{W}}(T,P) \le c\rho M.$$

Hence, since there are only finitely many $P_1,\ldots,P_N$ currents P as in (2) (N depends only on M, W, n, k, ρ), we have (1) as required.

Next note that (by 4.5(1),(2) and an argument as in 6.7(2) of Ch. 2) we can find a subsequence $\{T_{j'}\} \subset \{T_j\}$ and a sequence $\{W_i\}$, $W_i \subset\subset W_{i+1} \subset\subset U$, $\cup_{i=1}^{\infty} W_i = U$, such that $\sup_{j'\ge 1}\mathbb{M}(\partial(T_{j'}\llcorner W_i)) < \infty\ \forall i$. Thus from now on we can assume without loss of generality that $W \subset\subset U$ and

(4) $$\operatorname{spt} T_j \subset \overline{W}\ \ \forall j.$$

Then take any $\widetilde{W}$ such that $W \subset\subset \widetilde{W} \subset\subset U$ and apply (1) with $\varepsilon = 1, \frac{1}{2}, \frac{1}{4}$ etc. to extract a subsequence $\{T_{j_r}\}_{r=1,2,\ldots}$ from $\{T_j\}$ such that

$$d_{\widetilde{W}}(T_{j_{r+1}}, T_{j_r}) < 2^{-r}$$

and hence

(5) $$T_{j_{r+1}} - T_{j_r} = \partial R_r + S_r$$

where R_r, S_r are integer multiplicity,

$$\operatorname{spt} R_r \cup \operatorname{spt} S_r \subset \widetilde{W}$$
$$\mathbb{M}(R_r) + \mathbb{M}(S_r) \le \frac{1}{2^r}.$$

Therefore by 3.13 we can define integer multiplicity $R^{(\ell)}$, $S^{(\ell)}$ by the $\mathbb{M}$-absolutely convergent series

$$R^{(\ell)} = \textstyle\sum_{r=\ell}^{\infty} R_r,\ \ S^{(\ell)} = \sum_{r=\ell}^{\infty} S_r;$$

then

$$\mathbb{M}(R^{(\ell)}) + \mathbb{M}(S^{(\ell)}) \le 2^{-\ell+1}$$

and (from (5))

$$T - T_{j_\ell} = \partial R^{(\ell)} + S^{(\ell)}.$$

Thus we have a subsequence $\{T_{j_\ell}\}$ of $\{T_j\}$ such that $d_{\widetilde{W}}(T, T_{j_\ell}) \to 0$. Since we can thus extract a subsequence converging relative to $d_{\widetilde{W}}$ from any given subsequence of $\{T_j\}$, we then have $d_{\widetilde{W}}(T, T_j) \to 0$; since this can be repeated with $W = W_i$, $\widetilde{W} = W_{i+1}\ \forall i$ (W_i as above), the required result evidently follows. □

8 The Rectifiability and Compactness Theorems

Here we prove the important Rectifiability Theorem for currents T which, together with ∂T, have locally finite mass and which have the additional property that $\Theta^{*n}(\mu_t, x) > 0$ for μ_T-a.e. x. The main tool of the proof is the Structure Theorem 3.8 of Ch. 3. Having established the Rectifiability Theorem, we show (in 8.2, 8.3) that it is then straightforward to establish the Compactness 3.11. Although this proof of the Compactness Theorem has the advantage of being conceptually straightforward, it is rather lengthy if one takes into account the effort needed to prove the Structure Theorem. Recently B. Solomon [Sol82] showed that it is possible to prove the Compactness Theorem (and to develop the whole theory of integer multiplicity currents) without use of the structure theorem.

8.1 (Rectifiability Theorem.) *Suppose $T \in \mathcal{D}_n(U)$ is such that $\mathbb{M}_W(T)+\mathbb{M}_W(\partial T) < \infty$ for all $W \subset\subset U$, and $\Theta^{*n}(\mu_T, x) > 0$ for μ_T-a.e. $x \in U$. Then T is* rectifiable; *that is*

$$T = \underline{\tau}(M, \theta, \xi),^4$$

where M is countably n-rectifiable, $\mathcal{H}^n$-measurable, θ is a positive locally $\mathcal{H}^n$-integrable function on M, and $\xi(x)$ orients the approximate tangent space $T_x M$ of M for $\mathcal{H}^n$-a.e. $x \in M$ (i.e. $\xi|_x$ is a measurable function of x and $\xi|_x = \pm\tau_1\wedge\cdots\wedge\tau_n$, where $\tau_1, \ldots, \tau_n$ is an orthonormal basis for the approximate tangent space $T_x M$ of M, for $\mathcal{H}^n$-a.e. $x \in M$.)

Proof: First note that (by 3.3 of Ch. 1)

(1) $$\mathcal{H}^n\{x \in A : \Theta^{*n}(\mu_T, x) > t\} \le t^{-1}\mu_T(A) \le t^{-1}\mu_T(W)$$

for any subset $A \subset W$ and any open $W \subset\subset U$ and $t > 0$. In particular

(2) $$\mathcal{H}^n\{x \in U : \Theta^{*n}(\mu_T, x) = \infty\} = 0.$$

Notice that the same argument applies with ∂T in place of T in order to give

(3) $$\mathcal{H}^n\{x \in U : \Theta^{*n}(\mu_{\partial T}, x) = \infty\} = 0.$$

(Notice we could also conclude $\mathcal{H}^d\{x \in U : \Theta^{*d}(\mu_{\partial T}, x) = \infty\} = 0$ for any $d > 0$ by 3.3 of Ch. 1.)

Next notice that, because of (2), (3), and the fact that $\mathbb{M}_W(T) + \mathbb{M}_W(\partial T) < \infty\ \forall W \subset\subset U$, we can apply Remark 2.37(1) to deduce

(4) $$\mu_T\{x \in U : \Theta^{*n}(\mu_T, x) = \infty\} = 0,$$

and

(5) $$\mu_T\{x \in U : \Theta^{*n}(\mu_{\partial T}, x) = \infty\} = 0.$$

Now let

$$M = \{x \in U : \Theta^{*n}(\mu_T, x) > 0\}.$$

[4] The notation here is as for integer multiplicity rectifiable currents as in §3 of the present chapter. That is, $\underline{\tau}(M, \theta, \xi)(\omega) = \int_M \langle\xi, \omega\rangle\,\theta\, d\mathcal{H}^n$, although θ is not assumed to be integer-valued here.

Since we can write $M = \cup_{j=1}^{\infty} M_j$, where $M_j = \{x \in M : \Theta^{*n}(\mu_T, x) > 1/j\}$, we see from (1) that M is the countable union of sets of finite $\mathcal{H}^n$-measure. Suppose P is an $\mathcal{H}^n$-measurable purely unrectifiable subset of M. By the Structure Theorem 3.6 of Ch. 3 we have an orthogonal transformation Q of $\mathbb{R}^{n+\ell}$ such that $\mathcal{H}^n(p_\alpha(QP)) = 0$ for each $\alpha = (i_1, \ldots, i_n)$ with $1 \le i_1 < i_2 < \cdots < i_n \le n+\ell$ (where p_α is the projection $(x^1, \ldots, x^P) \mapsto (x^{i_1}, \ldots, x^{i_n})$). Then, by Remark 2.37(2), we conclude $\mu_T(P) = 0$ and hence

$$\text{(6)} \qquad \mathcal{H}^n(P) = 0 \ \ \forall \text{ purely unrectifiable } P \subset M$$

Then by Lemma 3.2 of Ch. 3 we conclude M is countably n-rectifiable. Thus we have proved that

$$\text{(7)} \qquad \mu_T = \mu_T \llcorner M,$$

with $M \subset U$ is countably n-rectifiable.

Since μ_T is absolutely continuous with respect to $\mathcal{H}^n$ (by Remark 2.37(1)), we can use the Radon-Nikodym Theorem 3.24 of Ch. 1 to conclude that

$$\text{(8)} \qquad \mu_T = \mathcal{H}^n \llcorner \theta,$$

where θ is a positive locally $\mathcal{H}^n$-integrable function on M and $\theta \equiv 0$ on $U \setminus M$. Then by the Riesz Representation Theorem 4.14 of Ch. 1 we have

$$\text{(9)} \qquad T(\omega) = \int_U \langle \xi, \omega \rangle \theta \, d\mathcal{H}^n,$$

for some $\mathcal{H}^n$-measurable, $\Lambda_n(\mathbb{R}^{n+\ell})$-valued function ξ, $|\xi| = 1$.

It thus remains only to prove that $\xi(x)$ orients $T_x M$ for $\mathcal{H}^n$-a.e. $x \in M$ (i.e. $\xi(x) = \pm \tau_1 \wedge \cdots \wedge \tau_n$ for $\mathcal{H}^n$-a.e. $x \in M$, where $\tau_1, \ldots, \tau_n$ is any orthonormal basis for the approximate tangent space $T_x M$ of M.) To see this, write $M = \cup_{j=0}^{\infty} M_j$, M_j pairwise disjoint, $\mathcal{H}^n(M_0) = 0$, $M_j \subset N_j$, N_j a C^1 submanifold of $\mathbb{R}^{n+\ell}$, $j \ge 1$. Now, by 3.6 of Ch. 1, if $j \ge 1$ we have, for $\mathcal{H}^n$-a.e. $x \in M_j$,

$$\text{(10)} \qquad \Theta^{*n}\big(\mu_T \llcorner ((N_j \setminus M_j) \cup (\cup_{k \ne j} M_k)), x\big) = 0.$$

Hence, writing as usual $\eta_{x,\lambda}(y) = \lambda^{-1}(y - x)$, we have for any $\omega \in \mathcal{D}^n(\mathbb{R}^{n+\ell})$ that, for all $x \in M_j$ such that (10) holds, and for λ small enough to ensure that $\operatorname{spt} \omega \subset \eta_{x,\lambda}(U)$,

$$\begin{aligned} \eta_{x,\lambda\#} T(\omega) &= T\left(\eta_{x,\lambda}{}^{\#} \omega\right) \\ &= \int_{N_j} \langle \xi, \eta_{x,\lambda}{}^{\#} \omega \rangle \, \theta \, d\mathcal{H}^n + \varepsilon(\lambda), \end{aligned}$$

where $\varepsilon(\lambda) \to 0$ as $\lambda \downarrow 0$. ($\varepsilon(\lambda)$ depending on x and ω.) That is, after the change of variable $z = \eta_{x,\lambda}(y)$ (i.e. $y = x + \lambda z$),

$$\eta_{x,\lambda\#} T(\omega) = \int_{\eta_{x,\lambda}(N_j)} \langle \xi(x + \lambda z, \omega(z)) \rangle \, \theta(x + \lambda z) \, d\mathcal{H}^n(z) + \varepsilon(\lambda)$$

$\mathcal{H}^n$-a.e. $x \in M_j$. Since N_j is C^1, this gives

$$\text{(11)} \qquad \lim_{\lambda \downarrow 0} \eta_{x,\lambda\#} T(\omega) = \theta(x) \int_L \langle \xi(x), \omega(z) \rangle \, d\mathcal{H}^n(z)$$

for $\mathcal{H}^n$-a.e. $x \in M_j$ (independent of ω), where L is the tangent space $T_x N_j$ of N_j at x. Thus (by definition of $T_x M$—see §2 of Ch. 3) we have (11) with $L = T_x M$ for $\mathcal{H}^n$-a.e. $x \in M_j$. On the other hand by (5) we have, provided $\operatorname{spt} \omega \subset B_R(0)$,

$$\text{(12)} \qquad \begin{aligned} \partial \eta_{x,\lambda\#} T(\omega) = \eta_{x,\lambda\#} \partial T(\omega) = \partial T\left(\eta_{x,\lambda}^{\#} \omega\right) &= \int_{B_{\lambda R}(x)} \left\langle \omega_{|\eta_{x,\lambda}(y)}, \eta_{x,\lambda\#} \overrightarrow{\partial T} \right\rangle d\mu_{\partial T} \\ &\le C|\omega| \lambda^{1-n} \mu_{\partial T}(B_{\lambda R}(x)) \to 0 \text{ as } \lambda \downarrow 0 \end{aligned}$$

for $\mathcal{H}^n$-a.e. $x \in M_j$ (independent of ω), because

$$\Theta^{*n}(\mu_{\partial T}, x) = \limsup_{\lambda \downarrow 0} \lambda^{-n} \mu_{\partial T}(B_\lambda(x)) < \infty \text{ for } \mathcal{H}^n\text{-a.e. } x \in M_j \text{ by (3).}$$

Thus, by (11) and (12), for $\mathcal{H}^n$-a.e. $x \in M$, we can find a sequence $\lambda_\ell \downarrow 0$ such that

$$\eta_{x,\lambda_\ell\#} T \rightharpoonup S_x, \ \partial S_x = 0,$$

where $S_x \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$ is defined by

$$\text{(13)} \qquad S_x(\omega) = \theta(x) \int_L \langle \xi(x), \omega(z) \rangle \, d\mathcal{H}^n(z),$$

$\omega \in \mathcal{D}^n(\mathbb{R}^{n+\ell})$, $L = T_x M$. We now claim that (13), taken together with the fact that $\partial S_x = 0$, implies that $\xi(x)$ orients L (i.e. $\xi_{|x} = \pm \tau_1 \wedge \cdots \wedge \tau_n$ with $\tau_1, \ldots, \tau_n$ an orthonormal basis for L). To see this, assume (without loss of generality) that $L = \mathbb{R}^n \times \{0\} \subset \mathbb{R}^{n+\ell}$ and select $\omega \in \mathcal{D}^{n-1}(\mathbb{R}^{n+\ell})$ so that $\omega(y) = y^j \varphi(y) dy^{i_1} \wedge \cdots \wedge dy^{i_{n-1}}$, where $y = (y^1, \ldots, y^{n+\ell})$, $j \ge n+1$, $\{i_1, \ldots, i_{n-1}\} \subset \{1, \ldots, n+\ell\}$, and $\varphi \in C_c^\infty(\mathbb{R}^{n+\ell})$. Then since $y_j \equiv 0$ on $\mathbb{R}^n \times \{0\}$ we deduce, from (13) and the fact that $\partial S_x = 0$,

$$\begin{aligned} 0 = \partial S_x(\omega) = S_x(d\omega) &= \theta(x) \int_L \varphi(y) \left\langle \xi(x), dy^j \wedge dy^{i_1} \wedge \cdots \wedge dy^{i_{n-1}} \right\rangle \\ &= \theta(x) \int_L \varphi(y) \xi(x) \cdot \left(e_j \wedge e_{i_1} \wedge \cdots \wedge e_{i_{n-1}}\right) d\mathcal{H}^n(y). \end{aligned}$$

That is, since $\varphi \in C_c^\infty(\mathbb{R}^{n+\ell})$ is arbitrary, we deduce that $\xi(x) \cdot (e_j \wedge e_{i_1} \wedge \cdots \wedge e_{i_{n-1}}) = 0$ whenever $j \ge n+1$ and $\{i_1, \ldots, i_{n-1}\} \subset \{1, \ldots, n+\ell\}$. Thus we must have (since $|\xi(x)| = 1$), $\xi(x) = \pm e_1 \wedge \cdots \wedge e_n$ as required. □

We can now give the proof of the Compactness 3.11. For convenience we first restate the theorem in a slightly weaker form. (See Remark 8.3(2) below for the proof of the previous version 3.11.)

8.2 Theorem. *Suppose $\{T_j\} \subset \mathcal{D}_n(U)$, suppose T_j, ∂T_j are integer multiplicity for each j,*

$$(\ddagger) \qquad \sup_{j\geq 1}\big(\mathbb{M}_W(T_j) + \mathbb{M}_W(\partial T_j)\big) < \infty \ \forall W \subset\subset U,$$

and suppose $T_j \rightharpoonup T \in \mathcal{D}_n(U)$. Then T is an integer multiplicity current.

8.3 Remarks (1) Note that the general case of the theorem follows from the special case when $U = \mathbb{R}^P$ and $\operatorname{spt} T_j \subset k$ for some fixed compact K; in fact if T_j are as in the theorem and if $\xi \in U$, then by 4.5(1), (2) and an argument like that in 6.7(2) of Ch. 2 we know that, for $\mathcal{L}^1$-a.e. $r > 0$, $\partial(T_{j'} \llcorner B_r(\xi))$ are integer multiplicity and 8.2 ($\ddagger$) holds with $T_{j'} \llcorner B_r(\xi)$ in place of T_j for some subsequence $\{j'\} \subset \{j\}$ (depending on r).

(2) The previous (formally slightly stronger) version 3.11 of the above theorem follows by using 6.3. (Note that the proof of 6.3 needed only the weaker version of the Compactness Theorem given above in 8.2; indeed, as in mentioned in 6.4, it used only the case $\partial T_j = 0$ of 3.11.

Proof of 8.2: We shall use induction on n with $U \subset \mathbb{R}^P$ (U, P fixed independent of n). First note that the theorem is trivial in case $n = 0$. Then assume $n \geq 1$ and suppose the theorem is true with $n-1$ in place of n.

By the above 8.3(1) we shall assume without loss of generality that $\operatorname{spt} T_j \subset K$ for some fixed compact K, and that $U = \mathbb{R}^P$. Furthermore, by 8.3(1) in combination with the inductive hypothesis, for each $\xi \in \mathbb{R}^P$ we have

$$(1) \qquad \partial(T \llcorner B_r(\xi)) \quad \text{is an integer multiplicity current}$$

(in $\mathcal{D}_{n-1}(\mathbb{R}^P)$) for $\mathcal{L}^1$-a.e. $r > 0$.

From the above assumptions $U = \mathbb{R}^P$, $\operatorname{spt} T_j \subset K$ we know that $0 \mathbin{⨳} \partial T - T$ zero boundary and is the weak limit of $0 \mathbin{⨳} \partial T_j - T_j$; since $0 \mathbin{⨳} \partial T$ is integer multiplicity (by the inductive hypothesis) we thus see that the general case of the theorem follows from the special case when $\partial T = 0$. We shall therefore henceforth also assume $\partial T = 0$.

Next, define (for $\xi \in \mathbb{R}^P$ fixed)

$$f(r) = \mathbb{M}(T \llcorner B_r(\xi)), \ r > 0.$$

By virtue of 4.9 we have (since $\partial T = 0$)

$$(2) \qquad \mathbb{M}(\partial(T \llcorner B_r(\xi))) \leq f'(r), \ \mathcal{L}^1\text{-a.e. } r > 0.$$

(Notice that $f'(r)$ exists a.e. $r > 0$ because $f(r)$ is increasing.) On the other hand if $\Theta^{*n}(\mu_T, \xi) < \eta$ ($\eta > 0$ a given constant), then $\limsup_{\rho\downarrow 0} \frac{f(\rho)}{\omega_n \rho^n} < \eta$, and hence for each $\delta > 0$ we can arrange

$$(3) \qquad \frac{d}{dr}\big(f^{1/n}(r)\big) \leq 2\omega_n^{1/n}\eta$$

for a set of $r \in (0, \delta)$ of positive $\mathcal{L}^1$-measure. (Because $\frac{1}{\delta}\int_0^\delta \frac{d}{dr}\big(f^{1/n}(r)\big)\,dr \leq \delta^{-1} f^{1/n}(\delta) \leq \omega_n^{1/n}\eta$ for all sufficiently small $\delta > 0$.)

Now by (1) and the Isoperimetric Theorem, we can find an integer multiplicity $S_r \in \mathcal{D}_n(\mathbb{R}^P)$ such that $\partial S_r = \partial(T \llcorner B_r(\xi))$ and

$$\begin{aligned}\mathbb{M}(S_r)^{\frac{n-1}{n}} &\leq c\mathbb{M}(\partial(T \llcorner B_r(\xi)))\\ &\leq c\eta\mathbb{M}(T \llcorner B_r(\xi))^{\frac{n-1}{n}} \quad \text{(by (2), (3))}\end{aligned}$$

for a set of r of positive $\mathcal{L}^1$-measure in $(0, \delta)$.[5] Since δ was arbitrary we then have both (1), (4) for a sequence of $r \downarrow 0$. But then (since we can repeat this for any ξ such that $\Theta^{*n}(\mu_T, \xi) < \eta$) if C is any compact subset of $\{x \in \mathbb{R}^P : \Theta^{*n}(\mu_T, x) < \eta\}$, by **??**(2) of Ch. 1 for each given $\rho > 0$ we get a pairwise disjoint family $B_j = B_{\rho_j}(\xi_j)$ of closed balls covering μ_T-almost all of C, with

$$\cup_j B_j \subset \{x : \operatorname{dist}(x, C) < \rho\}$$

and with

$$(4) \qquad \mathbb{M}(S_j^{(\rho)}) \leq c\eta\mathbb{M}(T \llcorner B_j)$$

for some integer multiplicity $S_j^{(\rho)}$ with

$$(5) \qquad \partial S_j^{(\rho)} = \partial(T \llcorner B_j).$$

Now because of (5) we have $S_j^{(\rho)} = T \llcorner B_j = \partial\big(\{\xi_j\} \mathbin{⨳} (S_j^{(\rho)} - T \llcorner B_j)\big)$, and hence (by 2.27, 2.30) we have for $\omega \in \mathcal{D}^n(\mathbb{R}^P)$

$$\begin{aligned}\big|(S_j^{(\rho)} - T \llcorner B_j)(\omega)\big| &\leq c\rho\mathbb{M}(S_j^{(\rho)} - T \llcorner B_j)|d\omega|\\ &\leq c\rho\mathbb{M}(T \llcorner B_j)|d\omega| \quad \text{(by (4))}.\end{aligned}$$

[5] In case $n = 1$, (1), (2), (3) (for $\eta < \frac{1}{4}$) imply $\partial(T \llcorner B_r(\xi)) = 0$, hence we get, in place of (4), $\mathbb{M}(S_r) \leq \mathbb{M}(T \llcorner B_r(\xi))$ trivially by taking $S_r = 0$.

Therefore we have $\sum_j \big(S_j^{(\rho)} - T \llcorner B_j\big) \rightharpoonup 0$ as $\rho \downarrow 0$, and hence

$$T + \textstyle\sum_j \big(S_j^{(\rho)} - T \llcorner B_j\big) \rightharpoonup T \tag{6}$$

as $\rho \downarrow 0$. However since the series $\sum_j S_j^{(\rho)}$ and $\sum_j T \llcorner B_j$ are $\mathbb{M}$-absolutely convergent (by (4) and the fact that the B_j are disjoint), we deduce that the left side in (6) can be written $T \llcorner (\mathbb{R}^P \setminus \cup_j B_j) + \sum_j S_j^{(\rho)}$ and hence (using (4) again, together with the lower semi-continuity of $\mathbb{M}_W$ (W open) under weak convergence)

$$\begin{aligned}\mu_T(\{x : \operatorname{dist}(x, C) < \rho\}) \le \mu_T(\{x &: \operatorname{dist}(x, C) < \rho\} \setminus C) \\ &+ c\eta\mu_T(\{x : \operatorname{dist}(x, C) < \rho\}).\end{aligned}$$

Choosing η such that $c\eta \le \frac{1}{2}$, this gives

$$\mu_T(\{x : \operatorname{dist}(x, C) < \rho\}) \le 2\mu_T(\{x : \operatorname{dist}(x, C) < \rho\} \setminus C)$$

Letting $\rho \downarrow 0$, we get $\mu_T(C) = 0$.

Thus we have shown that $\Theta^{*n}(\mu_T, x) > 0$ for μ_T-a.e. $x \in \mathbb{R}^P$. We can therefore apply 8.1 in order to conclude that $T = \underline{\tau}(M, \theta, \xi)$ as in 8.1. It thus remains only to prove that θ is integer-valued. This is achieved as follows:

First note that for $\mathcal{L}^n$-a.e. $x \in M$ we have (Cf. the argument leading to (11) in the proof of Theorem 8.1)

$$\eta_{x,\lambda\#}T \rightharpoonup \theta(x)[\![T_x M]\!] \quad \text{as} \quad \lambda \downarrow 0, \tag{7}$$

where $[\![T_x]\!]$ is oriented by $\xi(x)$. Assuming without loss of generality that $T_x M = \mathbb{R}^n \times \{0\} \subset \mathbb{R}^P$ and setting $d(y) = \operatorname{dist}(y, \mathbb{R}^n \times \{0\})$, by 4.5(1) we can find a sequence $\lambda_j \downarrow 0$ and a $\rho > 0$ such that $\langle \eta_{x,\lambda_j\#}T, d, \rho\rangle$ is integer multiplicity with

$$\mathbb{M}_\Omega\big(\langle \eta_{x,\lambda_j\#}T, d, \rho\rangle\big) \le c \quad \text{(independent of } j) \tag{8}$$

where $\Omega = B_1^n(0) \times \mathbb{R}^{P-n} \subset \mathbb{R}^P$. Next, we choose $\{j'\} \subset \{j\}$ and $\rho > 0$ so that $\eta_{x,\lambda_j\#}T_{j'} \rightharpoonup \theta(x)[\![T_x M]\!]$ (which is possible by (7) and the fact that $T_j \rightharpoonup T$), and so that (8) remains valid with $T_{j'}$ instead of T (which is justified by 4.5(1) and a selection arguments in 6.7(2) of Ch. 2). Then by 4.5(2) we have $S_j \equiv \big(\eta_{x,\lambda_j\#}T_j\big) \llcorner \{y : d(y) < \rho\}$ is such that

$$\sup_{j\ge 1}\big(\mathbb{M}_\Omega(S_j) + \mathbb{M}_\Omega(\partial S_j)\big) < \infty \tag{9}$$

with $\Omega = B_1^n(0) \times \mathbb{R}^{P-n} \subset \mathbb{R}^P$. Now let p denote the restriction to Ω of the orthogonal projection of $\mathbb{R}^P$ onto $\mathbb{R}^n$; and let $\widetilde{S}_j$ be the current in $\mathcal{D}_n(\Omega)$ obtained

by setting $\widetilde{S}_j(\omega) = S_j(\widetilde{\omega})$, $\omega \in \mathcal{D}^n(\Omega)$, $\widetilde{\omega} \in \mathcal{D}^n(\mathbb{R}^P)$ such that $\widetilde{\omega} = \omega$ in Ω and $\widetilde{\omega} \equiv 0$ on $\mathbb{R}^P \setminus \Omega$. Then we have $p_\#\widetilde{S}_j \in \mathcal{D}_n(B_1^n(0))$, and hence, by 2.35 and (9) above,

$$p_\#\widetilde{S}_j(\omega) = \int_{B_1^n(0)} a\theta_j \, d\mathcal{L}^n, \ \omega = a\, dx^1 \wedge \cdots \wedge dx^n, \ a \in C_c^\infty(\mathbb{R}^n),$$

for some integer-valued $BV_{\mathrm{loc}}(B_1^n(0))$ function θ_j with

$$\begin{cases} \mathbb{M}_{\breve{B}_1^n(0)}(p_\#\widetilde{S}_j) = \int_{B_1^n(0)} |\theta_j| \, d\mathcal{L}^n \\ \mathbb{M}_{\breve{B}_1^n(0)}(\partial p_\#\widetilde{S}_j) = \int_{B_1^n(0)} |D\theta_j|. \end{cases} \tag{10}$$

Then by (9), (10) we deduce $\int_{B_1^n(0)} |D\theta_j| + \int_{B_1^n(0)} |\theta_j| \, d\mathcal{L}^n \le c$, c independent of j, and hence by the Compactness Theorem 2.6 of Ch. 2 we know θ_j converges strongly in L^1 in $B_1^n(0)$ to an integer-valued BV function θ_*. On the other hand $S_j \rightharpoonup \theta(x)[\![\mathbb{R}^n \times \{0\}]\!]$ by (7), and hence $p_\#\widetilde{S}_j \rightharpoonup \theta(x)p_\#[\![\mathbb{R}^n \times \{0\}]\!] = \theta(x)[\![\mathbb{R}^n]\!]$ in $B_1^n(0)$. We thus deduce that $\theta_* \equiv \theta(x)$ in $B_1^n(0)$; thus $\theta(x) \in \mathbb{Z}$ as required. $\square$

Chapter 7

Area Minimizing Currents

1 Basic Concepts

Suppose A is any subset of $\mathbb{R}^{n+\ell}$, $A \subset U$, U open in $\mathbb{R}^{n+\ell}$, and $T \in \mathcal{D}_n(U)$ an integer multiplicity current.

1.1 Definition: We say that T is minimizing in A if

$$\mathbb{M}_W(T) \leq \mathbb{M}_W(S)$$

whenever $W \subset\subset U$, $\partial S = \partial T$ (in U) and $\operatorname{spt}(S - T)$ is a compact subset of $A \cap W$.

There are two especially important cases (in fact the only cases we are interested in here) of this definition:

(1) when $A = U$

(2) when $A = N \cap U$, N an $(n+\ell)$-dimensional embedded C^2 submanifold of $\mathbb{R}^{n+\ell}$ (in the sense of §4 of Ch. 2).

Corresponding to the current $T = \underline{\tau}(M, \theta, \xi) \in \mathcal{D}_n(U)$ we have the integer multiplicity varifold $V = \underline{v}(M, \theta)$. As one would expect, V is stationary in U if T is minimizing in U and $\partial T = 0$:

1.2 Lemma. *Suppose T is minimizing in $N \cap U$, where N is an $(n+\ell)$-dimensional*

C^2 submanifold of $\mathbb{R}^{n+\ell}$ ($\ell \le k$, so $N = \mathbb{R}^{n+\ell}$ is an important special case) and suppose $\partial T = 0$ in U. Then V is stationary in $N \cap U$ in the sense of 2.6 *of* Ch. 4*, so that in particular V has locally bounded generalized mean curvature in U (in the sense of* 3.14 *of* Ch. 4*).*

In fact V is minimizing in $N \cap U$ in the sense that

(‡) $$\mathbb{M}_W(V) \le \mathbb{M}_W(\varphi_\# V),$$

whenever $W \subset\subset U$ and φ is a diffeomorphism of U such that $\varphi(N \cap U) \subset N \cap U$ and $\varphi\big|U \setminus K = 1_{U\setminus K}$ for some compact $K \subset W \cap N$.

1.3 Remark: In view of 1.2 (together with the fact that $\theta \ge 1$) we can represent $T = \underline{\tau}(M_*, \theta_*, \xi)$ where M_* is a relatively *closed* countably n-rectifiable subset of U, and θ_* is an upper semi-continuous function on M_* with $\theta_* \ge 1$ everywhere on M_* (and θ_* integer-valued $\mathcal{H}^n$-a.e. on M_*).

Proof of 1.2: Evidently (in view of the discussion of §2 of Ch. 4) the first claim in 1.2 follows from 1.2 (‡) (by taking $\varphi = \varphi_t$ in 1.2 (‡), φ_t is in 2.1 of Ch. 4 with $U \cap N$ in place of U).

To prove 1.2 (‡) we first note that, for any W, φ as in the statement of the theorem,

(1) $$\mathbb{M}_W(\varphi_\# V) = \mathbb{M}_W(\varphi_\# T)$$

by 3.2(3) of Ch. 6. Also, since $\partial T = 0$ (in U), we have

(2) $$\partial \varphi_\# T = \varphi_\# \partial T = 0.$$

Finally,

(3) $$\operatorname{spt}(T - \varphi_\# T) \subset K \subset W.$$

By virtue of (2), (3) we are able to use the inequality of 1.1 with $S = \varphi_\# T$. This gives 1.2 (‡) as required by virtue of (1). □

We conclude this section with the following useful decomposition lemma:

1.4 Lemma. *Suppose $T_1, T_2 \in \mathcal{D}_n(U)$ are integer multiplicity and suppose $T_1 + T_2$ is minimizing in A, $A \subset U$, and*

$$\mathbb{M}_W(T_1 + T_2) = \mathbb{M}_W(T_1) + \mathbb{M}_W(T_2)$$

for each $W \subset\subset U$. Then T_1, T_2 are both minimizing in A.

Proof: Let $X \in \mathcal{D}_n(U)$ be integer multiplicity with $\operatorname{spt} X \subset K$, K a compact subset of $A \cap W$, and with $\partial X = 0$. Because $T_1 + T_2$ is minimizing in A we have (by 1.1)

$$\mathbb{M}_W(T_1 + T_2 + X) \ge \mathbb{M}_W(T_1 + T_2).$$

However since $\mathbb{M}_W(T_1 + T_2) = \mathbb{M}_W(T_1) + \mathbb{M}_W(T_2)$, and $\mathbb{M}(T_1 + T_2 + X) \le \mathbb{M}_W(T_1 + X) + \mathbb{M}_W(T_2)$, this gives

$$\mathbb{M}_W(T_1) \le \mathbb{M}_W(T_1 + X).$$

In view of the arbitrariness of X, this establishes that T_1 is minimizing in $A \cap W$ (in accordance with 1.1). Interchanging T_1, T_2 in the above argument, we likewise deduce that T_2 is minimizing in $A \cap W$. □

2 Existence and Compactness Results

We begin with a result which establishes the rich abundance of area minimizing currents in Euclidean space.

2.1 Lemma. *Let $S \in \mathcal{D}_{n-1}(\mathbb{R}^{n+\ell})$ be integer multiplicity with* spt S *compact and $\partial S = 0$. Then there is an integer multiplicity current $T \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$ such that* spt T *is compact and $\mathbb{M}(T) \le \mathbb{M}(R)$ for each integer multiplicity $R \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$ with* spt R *compact and $\partial R = S$.*

2.2 Remarks: (1) Of course T is minimizing in $\mathbb{R}^{n+\ell}$ in the sense of 1.1.

(2) By virtue of 1.2 and the convex hull property (Theorem 6.2 of Ch. 4) we have automatically that spt $T \subset$ convex hull of spt S.

(3) $\mathbb{M}(T)^{\frac{n-1}{n}} \le c\mathbb{M}(S)$ by virtue of the Isoperimetric Inequality 6.1 of Ch. 6.

Proof of 2.1: Let

$$\mathcal{I}_S = \left\{ R \in \mathcal{D}_n(\mathbb{R}^{n+\ell}) : R \text{ is integer multiplicity, } \operatorname{spt} R \text{ compact, } \partial R = S \right\}.$$

Evidently $\mathcal{I}_S \ne \emptyset$. (e.g. $0 \mathbin{⨳} S \in \mathcal{I}_S$.) Take any sequence $\{R_q\} \subset \mathcal{I}_S$ with

(1) $$\lim_{q\to\infty} \mathbb{M}(R_q) = \inf_{R\in\mathcal{I}_S} \mathbb{M}(R),$$

let $B_R(0)$ by any ball in $\mathbb{R}^{n+\ell}$ such that spt $S \subset B_R(0)$, and let $f : \mathbb{R}^{n+\ell} \to B_R(0)$ be the nearest point (radial) retract of $\mathbb{R}^{n+\ell}$ onto $B_R(0)$. Then Lip $f = 1$ and hence

(2) $$\mathbb{M}(f_\# R_q) \le \mathbb{M}(R_q).$$

On the other hand $\partial f_\# R_q = f_\# \partial R_q = f_\# S = S$, because $f\big|B_R(0) = 1_{B_R(0)}$ and spt $S \subset B_R(0)$. Thus $f_\# R_q \subset \mathcal{I}_S$ and by (1), (2) we have

(3) $$\lim_{q\to\infty} \mathbb{M}(f_\# R_q) = \inf_{R\in\mathcal{I}_S} \mathbb{M}(R).$$

Now by the Compactness Theorem 3.11 of Ch. 6 there is a subsequence $\{q'\} \subset \{q\}$ and an integer multiplicity current $T \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$ such that $f_\# R_{q'} \rightharpoonup T$ and (by (3) and lower semi-continuity of mass with respect to weak convergence)

$$\mathbb{M}(T) \leq \inf_{R \in \mathcal{I}_S} \mathbb{M}(R). \tag{4}$$

However $\operatorname{spt} T \subset B_R(0)$ and $\partial T = \lim \partial f_\# R_{q'} = \lim f_\# \partial R_{q'} = S$, so that $T \in \mathcal{I}_S$, and the lemma is established (by (4)). □

The proof of the following lemma is similar to that of 2.1 (and again based on 3.11 of Ch. 6), and its proof is left to the reader.

Lemma. *Suppose N is an $(n+\ell)$-dimensional compact C^1 submanifold embedded in $\mathbb{R}^{n+\ell}$ and suppose $R_1 \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$ is given such that $\partial R_1 = 0$, $\operatorname{spt} R_1 \subset N$ and*

$$\mathcal{I}_{R_1} = \Big\{ R \in \mathcal{D}_n(\mathbb{R}^{n+\ell}) : R - R_1 = \partial S \text{ for some integer multiplicity } S \in \mathcal{D}_{n+1}(\mathbb{R}^{n+\ell}) \text{ with } \operatorname{spt} S \subset N \Big\} \neq \emptyset.$$

Then there is $T \in \mathcal{I}_{R_1}$ such that

$$\mathbb{M}(T) = \inf_{R \in \mathcal{I}_{R_1}} \mathbb{M}(R).$$

2.3 Remarks: (1) $R - R_1 = \partial S$ with S integer multiplicity and $\operatorname{spt} S \subset N$ means that R, R_1 represents homologous cycles in the n-th singular homology class (with integer coefficients) of N (See [Fed69] or [FF60] for discussion.)

(2) It is quite easy to see that T is *locally* minimizing in N; thus for each $\xi \in \operatorname{spt} T$ there is a neighborhood U of ξ such that T is minimizing in $N \cap U$.

We conclude this section with the following important compactness theorem for minimizing currents:

2.4 Theorem. *Suppose $\{T_j\}$ is a sequence of minimizing currents in U with*

$$\sup_{j \geq 1} (\mathbb{M}_W(T_j) + \mathbb{M}_W(\partial T_j)) < \infty \text{ for each } W \subset\subset U,$$

and suppose $T_j \rightharpoonup T \in \mathcal{D}_n(U)$. Then T is minimizing in U and $\mu_{T_j} \to \mu_T$ (in the usual sense of Radon measures in U).

2.5 Remarks: (1) Note that $\mu_{T_j} \to \mu_T$ means the corresponding sequence of varifolds converge in the measure theoretic sense of §1 of Ch. 4 to the varifold associated with T. (T is automatically integer multiplicity by 3.11 of Ch. 6.)

(2) If the hypotheses are as in the theorem, except that $\operatorname{spt} T_j \subset N_j \subset U$ and T_j is minimizing in N_j, $\{N_j\}$ a sequence of C^1 embedded $(n+\ell)$-dimensional submanifolds of $\mathbb{R}^{n+\ell}$ converging in the C^1 sense to N, $N \subset U$ an embedded $(n+\ell)$-dimensional C^1 submanifold of $\mathbb{R}^{n+\ell}$,[1] then T minimizes in N (and we still have $\mu_{T_j} \to \mu_T$ in the sense of Radon measures in U). We leave this modification of 2.4 to the reader. (It is easily checked by using suitable local representations for the N_j and by obvious modifications of the proof of 2.4 given below.)

Proof of 2.4: Let $K \subset U$ be an arbitrary compact set and choose a smooth $\varphi : U \to [0,1]$ such that $\varphi \equiv 1$ in some neighborhood of K, and $\operatorname{spt} \varphi \subset \{x \in U : \operatorname{dist}(x, K) < \varepsilon\}$, where $0 < \varepsilon < \operatorname{dist}(K, \partial U)$ is arbitrary. For $0 < \lambda < 1$, let

$$W_\lambda = \{x \in U : \varphi(x) > \lambda\}.$$

Then

$$K \subset W_\lambda \subset\subset U \tag{1}$$

for each λ, $0 \leq \lambda < 1$.

By virtue of 7.2 of Ch. 6 we know that $d_W(T_j, T) \to 0$ for each $W \subset\subset U$, hence in particular we have

$$T - T_j = \partial R_j + S_j, \; \mathbb{M}_{W_0}(R_j) + \mathbb{M}_{W_0}(S_j) \to 0 \tag{2}$$

($W_0 = \{x \in U : \varphi(x) > 0\}$).

By the slicing theory (and in particular by 4.5 of Ch. 6) we can choose $0 < \alpha < 1$ and a subsequence $\{j'\} \subset \{j\}$ (subsequently denoted simply by $\{j\}$) such that

$$\partial(R_j \llcorner W_\alpha) = (\partial R_j) \llcorner W_\alpha + P_j \tag{3}$$

where $\operatorname{spt} P_j \subset \partial W_\alpha$, P_j is integer multiplicity, and

$$\mathbb{M}(P_j) \to 0. \tag{4}$$

We can also of course choose α to be such that

$$\mathbb{M}(T_j \llcorner \partial W_\alpha) = 0 \; \forall j \text{ and } \mathbb{M}(T \llcorner \partial W_\alpha) = 0. \tag{5}$$

Thus, combining (2), (3), (4) we have

$$T \llcorner W_\alpha = T_j \llcorner W_\alpha + \partial \widetilde{R}_j + \widetilde{S}_j \tag{6}$$

[1] Thus $\exists \psi_j : U \to U$, $\psi_j | N_j$ in a diffeomorphism onto N, and $\psi_j \to 1_U$ locally in U with respect to the C^1 metric.

with $\widetilde{R}_j, \widetilde{S}_j$ integer multiplicity ($\widetilde{R}_j = R_j \llcorner W_\alpha$, $\widetilde{S}_j = S_j \llcorner W_\alpha + P_j$) with

$$\mathbb{M}(\widetilde{R}_j) + \mathbb{M}(\widetilde{S}_j) \to 0. \tag{7}$$

Now let $X \in \mathcal{D}_n(U)$ be any integer multiplicity current with $\partial X = 0$ and $\operatorname{spt} X \subset K$. We want to prove

$$\mathbb{M}_{W_\alpha}(T) \le \mathbb{M}_{W_\alpha}(T + X). \tag{8}$$

(In view of the arbitrariness of K, X this will evidently establish the fact that T is minimizing in U.)

By (6), we have

$$\begin{aligned}\mathbb{M}_{W_\alpha}(T + X) &= \mathbb{M}_{W_\alpha}(T_j + X + \partial\widetilde{R}_j + \widetilde{S}_j) \\ &\ge \mathbb{M}_{W_\alpha}(T_j + X + \partial\widetilde{R}_j) - \mathbb{M}(\widetilde{S}_j).\end{aligned}$$

Now since T_j is minimizing and $\partial(X + \partial\widetilde{R}_j) = 0$ with $\operatorname{spt}(X + \partial\widetilde{R}_j) \subset \overline{W}_\alpha$, we have

$$\mathbb{M}_{W_\lambda}(T_j + X + \partial\widetilde{R}_j) \ge \mathbb{M}_{W_\lambda}(T_j) \tag{9}$$

for $\lambda > \alpha$. But by (3) we have $\mathbb{M}(\partial\widetilde{R}_j \llcorner \partial W_\alpha) = \mathbb{M}(P_j) \to 0$, and by (5) $\mathbb{M}(T_j \llcorner \partial W_\alpha) = 0$, $(T \llcorner \partial W_\alpha) = 0$. Hence letting $\lambda \downarrow 0$ in (9) we get

$$\mathbb{M}_{W_\alpha}(T_j + X + \partial\widetilde{R}_j) \ge \mathbb{M}_{W_\alpha}(T_j) - \mathbb{M}(P_j), \tag{10}$$

and therefore from (9) we obtain

$$\mathbb{M}_{W_\alpha}(T + X) \ge \mathbb{M}_{W_\alpha}(T_j) - \varepsilon_j, \quad \varepsilon_j \downarrow 0. \tag{11}$$

In particular, setting $X = 0$, we have

$$\mathbb{M}_{W_\alpha}(T) \ge \mathbb{M}_{W_\alpha}(T_j) - \varepsilon_j, \quad \varepsilon_j \downarrow 0. \tag{12}$$

Using the lower semi-continuity of mass with respect to weak convergence in (11), we then have (8) as required.

It thus remains only to prove that $\mu_{T_j} \to \mu_T$ in the sense of Radon measures in U. First note that by (12) we have

$$\limsup \mathbb{M}_{W_\alpha}(T_j) \le \mathbb{M}_{W_\alpha}(T),$$

so that (since $K \subset W_\alpha \subset \{x : \operatorname{dist}(x, K) < \varepsilon\}$ by construction)

$$\limsup \mu_{T_j}(K) \le \mathbb{M}_{\{x:\operatorname{dist}(x,K)<\varepsilon\}}(T).$$

Hence, letting $\varepsilon \downarrow 0$

$$\limsup \mu_{T_j}(K) \le \mu_T(K). \tag{13}$$

(We actually only proved this for some subsequence, but we can repeat the argument fora subsequence of any given subsequence, hence it holds for the original sequence $\{T_j\}$.)

By the lower semi-continuity of mass with respect to weak convergence we have

$$\mu_T(W) \le \liminf \mu_{T_j}(W) \ \forall \text{ open } W \subset\subset U. \tag{14}$$

Since (13), (14) hold for arbitrary compact K and open $W \subset U$, it now easily follows (by a standard approximation argument) that $\int f\, d\mu_{T_j} \to \int f\, d\mu_T$ for each continuous f with compact support in U, as required. □

3 Tangent Cones and Densities

In this section we prove the basic results concerning tangent cones and densities of area minimizing currents. All results depend on the fact that (by virtue of 1.2 the varifold associated with a minimizing current is stationary. This enables us to bring into play the important monotonicity results of §4 of Ch.4.

Subsequently we take N to be a smooth (at least C^2) embedded $(n+\ell)$-dimensional submanifold of $\mathbb{R}^{n+\ell}$ ($\ell \le k$), U open in $\mathbb{R}^{n+\ell}$ and $(\overline{N} \setminus N) \cap U = \emptyset$. Notice that an important case is when $N = U$ (when $\ell = k$).

3.1 Theorem. *Suppose $T \in \mathcal{D}_n(U)$ is minimizing in $U \cap N$, $\operatorname{spt} T \subset U \cap N$, and $x \in \operatorname{spt} T \setminus \operatorname{spt} \partial T$. Then*

(1) *$\Theta^n(\mu_T, x)$ exists everywhere in U and $\Theta^n(\mu_T, \cdot)$ is upper semi-continuous in U;*

(2) *For each $x \in \operatorname{spt} T$ and each sequence $\{\lambda_j\} \downarrow 0$, there is a subsequence $\{\lambda_{j'}\}$ such that $\eta_{x,\lambda_{j'}\#} T \rightharpoonup C$ and $\mu_{\eta_{x,\lambda_{j'}\#}T} \to \mu_C$ in $\mathbb{R}^{n+\ell}$, where $C \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$ is integer multiplicity and minimizing in $\mathbb{R}^{n+\ell}$, $\eta_{0,\lambda\#} C = C\ \forall \lambda > 0$, and $\Theta^n(\mu_C, 0) = \Theta^n(\mu_T, x)$.*

3.2 Remarks: (1) If C is as in 3.1(2) above, we say that C is a *tangent cone* for T at x. If $\operatorname{spt} C$ is an n-dimensional subspace P. Notice that since C is integer multiplicity and $\partial C = 0$, it then follows from 2.34 of Ch. 6 that, assuming we chose an appropriate (constant) orientation for P, $C = m[\![P]\!]$ for some $m \in \{1, 2, \ldots\}$. In this case we call C a *multiplicity m tangent plane for T at x.*

(2) Notice that is *not* clear whether or not there is an *unique* tangent cone for T at x; thus it is an open question whether or not C depends on the particular sequence $\{\lambda_j\}$ or subsequence $\{\lambda_{j'}\}$ use in its definition. The work of [Sim83] shows that if C is a tangent cone of T at x such that $\Theta^n(\mu_C, x) = 1$ for *all* $x \in \operatorname{spt} C \setminus \{0\}$, then C is the unique tangent cone for T at x, and hence $\eta_{x,\lambda\#}T \rightharpoonup C$ as $\lambda \downarrow 0$. Also B. White [Whi82] has shown in case $n = 2$ that C is always unique with spt C consisting of a union of 2-planes.

Proof of 3.1: By virtue of 1.2 we can apply the monotonicity formula of 4.7 of Ch.4 (with $\alpha = 1$) and 4.9 of Ch.4 in order to deduce that $\Theta^n(\mu_T, x)$ exists for every $x \in U$ and is an upper semi-continuous function of x in U.

Thus in particular

$$(1) \qquad (\omega_n R^n)^{-1}\mathbb{M}_{\breve{B}_R(0)}(\eta_{x,\lambda_j\#}T) = (\omega_n\lambda_j^n R^n)^{-1}\mathbb{M}_{\breve{B}_{\lambda_j R}(x)}(T) \to \Theta^n(\mu_T, x)$$

for each $R > 0$, and hence $\sup_j \mathbb{M}_{\breve{B}_R(0)}(\eta_{x,\lambda_j\#}T) < \infty$ for each $R > 0$, while $\partial\eta_{x,\lambda_j\#}T = 0$ in $\breve{B}_R(0)$ for sufficiently large j (because $x \notin \operatorname{spt}\partial T$), so we can apply the compactness theorem 2.4 to give a subsequence j' such $\eta_{x,\lambda_{j'}\#}T \rightharpoonup C$ in $\mathbb{R}^{n+\ell}$ with C integer multiplicity minimizing[2], so

$$(2) \qquad C = \tau(\operatorname{spt} C, \xi, \Theta^n(\mu_C, \cdot)),$$

$\mu_{\eta_{x,\lambda_{j'}\#}T} \to \mu_C$ in $\mathbb{R}^{n+\ell}$, and (by Lemma 1.2) the rectifiable varifold

$$(3) \qquad V_C = \underline{v}(\operatorname{spt} C, \Theta^n(\mu_C, \cdot))$$

is stationary in $\mathbb{R}^{n+\ell}$. In particular for any $\rho > 0$ with $\mu_C(\partial B_\rho(0)) = 0$ (which is true except for at most a countable set of ρ) we have

$$(4) \qquad \mu_{\eta_{x,\lambda_{j'}\#}T}(B_\rho(0)) \to \mu_C(B_\rho(0)),$$

and together with (1) gives $(\omega_n\rho^n)^{-1}\mu_C(B_\rho(0)) = \Theta^n(\mu_T, x)$ for each $\rho > 0$. Then by the monotonicity formula 3.6 of Ch.4, applied to the stationary varifold V_C of (3), we have $D^\perp r = 0$ μ_C-a.e., where $r = |x|$, and $D^\perp r$ is orthogonal projection of $Dr = r^{-1}x$ onto the normal space $(T_x \operatorname{spt} C)^\perp$. That is $x \in T_x \operatorname{spt} C$ for μ_C-a.e. x, so in particular $x \wedge \vec{C} = 0$ μ_C-a.e. and hence we can apply Lemma 2.33 to deduce that C is a cone. □

[2]See Remark 2.5; notice this establishes first that C is minimizing only in the $(n+\ell)$-dimensional subspace $T_x N \subset \mathbb{R}^{n+\ell}$. However since orthogonal projection of $\mathbb{R}^{n+\ell}$ onto $T_x N$ does not increase area, and since spt $C \subset T_x N$, it then follows that C is area minimizing in $\mathbb{R}^{n+\ell}$ as claimed.

3.3 Theorem.[3] *Suppose $T \in \mathcal{D}_n(U)$ is minimizing in $U \cap N$, spt $T \subset U \cap N$, and $\partial T = 0$ (in U). Then*

(1) $\Theta^n(\mu_T, x) \in \mathbb{Z}$ for all $x \in U \setminus E$, where $\mathcal{H}^{n-3+\alpha}(E) = 0\ \forall \alpha > 0$;

(2) There is a set $F \subset E$ (E as in (1)) with $\mathcal{H}^{n-2+\alpha}(F) = 0\ \forall\alpha > 0$ and such that for each $x \in \operatorname{spt} T \setminus F$ there is a tangent plane (see 3.2(1) above for terminology) for T at x.

Note: We do not claim E, R are *closed.*

The proof of both parts is based on the abstract dimension reducing argument of Appendix A. In order to apply this in the context of currents we need the observation of the following remark.

3.4 Remark: Given an integer multiplicity current $S \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$, there is an associated function $\varphi_S = (\varphi_S^0, \varphi_S^1, \ldots, \varphi_S^N) : \mathbb{R}^{n+\ell} \to \mathbb{R}^{N+1}$, where $N = \binom{n+\ell}{n}$, such that (writing $\theta_S(x) = \Theta^{*n}(\mu_S, x)$)

$$\varphi_S^0(x) = \theta_S(x),\ \varphi_S^j(x) = \theta_S(x)\xi_S^j(x),\ j = 1, \ldots, N,$$

where $\xi_S^j(x)$ in the j-th component of the orientation $\vec{S}(x)$ relative to the usual orthonormal basis $e_{i_1} \wedge \cdots \wedge e_{i_n}$, $1 \le i_1 < i_2 < \cdots < i_n \le n+\ell$ for $\Lambda_n(\mathbb{R}^{n+\ell})$ (ordered in any convenient manner). Evidently, for any $x \in \mathbb{R}^{n+\ell}$,

$$\varphi_S(x + \lambda y) = \varphi_{\eta_{x,\lambda\#}S}(y),\ y \in \mathbb{R}^{n+\ell},$$

and, given a sequence $\{S_i\} \subset \mathcal{D}_n(I + \mathbb{R}^{n+\ell})$ of such integer multiplicity currents, we trivially have

$$\varphi_{S_i}^j\, d\mathcal{H}^n \to \varphi_S^j\, d\mathcal{H}^n\ \forall j \in \{1, \ldots, N\} \iff S_i \rightharpoonup S$$

and

$$\varphi_{S_i}^0\, d\mathcal{H}^n \to \varphi_S^0\, d\mathcal{H}^n \iff \mu_{S_i} \to \mu_S$$

(where $\psi_i\, d\mathcal{H}^n \to \psi\, d\mathcal{H}^n$ means $\int f\psi_i\, d\mathcal{H}^n \to \int f\psi\, d\mathcal{H}^n\ \forall f \in C_c(\mathbb{R}^{n+\ell})$).

We shall also need the following simple lemma, the proof of which is left to the reader.

3.5 Lemma. *Suppose S is minimizing in $\mathbb{R}^{n+\ell}$, $\partial S = 0$, and*

$$\eta_{x,1\#}S = S\ \forall x \in \mathbb{R}^m \times \{0\} \subset \mathbb{R}^{n+\ell}$$

[3] Cf. Almgren [Alm84]

for some positive integer $m < n$. (Recall $\eta_{x,1} : y \mapsto y - x$, $y \in \mathbb{R}^{n+\ell}$.) Then

$$S = [\![\mathbb{R}^m]\!] \times S_0,$$

where $\partial S_0 = 0$ and S_0 is minimizing in $\mathbb{R}^{n+\ell-m}$.

Furthermore if S is a cone (i.e. $\eta_{0,\lambda\#} S = S$ for each $\lambda > 0$), then so is S_0.

Proof of 3.3(1): For each positive integer m and $\beta \in (0, \frac{1}{2})$ let

(1) $$U_{m,\beta} = \{x \in U : \Theta^n(\mu_{T,x}) < m - \beta\}.$$

Now T is minimizing in $U \cap N$, so by the monotonicity of 4.7 of Ch. 4 (which can be applied by virtue of 1.2) we have, firstly, that $U_{m,\beta}$ is open, and secondly that for each $x \in U_{m,\beta}$, there is some ball $B_{2\rho}(x) \subset U_{m,\beta}$ such that

(2) $$\frac{\mu_T(B_\sigma(y))}{\omega_n \sigma^n} \le m - \frac{\beta}{2} \ \forall \sigma < \rho, \ y \in B_\rho(x).$$

We ultimately want to prove

(3) $$\mathcal{H}^{n-3+\alpha}\left(\cup_{m=1}^\infty \{x \in U_{m,\beta} : m - 1 + \beta < \Theta^n(\mu_T, x) < m - \beta\}\right) = 0$$

for each sufficiently small $\alpha, \beta > 0$ and, in view of (2), by a rescaling and translation it will evidently suffice to assume

(4) $$B_2(0) = U, \ \frac{\mu_T(B_\sigma(y))}{\omega_n \sigma^n} \le m - \beta \ \forall \sigma < 1, \ y \in B_1(0),$$

and then prove

(5) $$\mathcal{H}^{n-3+\alpha}\{x \in B_1(0) : \Theta^n(\mu_T, x) \ge m - 1 + \beta\} = 0.$$

We consider the set $\mathcal{T}$ of weak limit points of sequences $S_i = \eta_{x_i,\lambda_i\#} T$ where $|x_i| < 1 - \lambda_i$, $0 < \lambda_i < 1$, with $\lim x_i \in B_1(0)$ and $\lim \lambda_i = \lambda \ge 0$ both existing. For any such sequence S_i we have (by (4))

(6) $$\limsup \mathbb{M}_W(S_i) < \infty$$

for each $W \subset\subset \eta_{x,\lambda}(U)$ in case $\lambda > 0$, an for each $W \subset\subset \mathbb{R}^{n+\ell}$ in case $\lambda = 0$. Hence we can apply the Compactness 2.4 to conclude that each element S of $\mathcal{T}$ is integer multiplicity and

(7) $$S \text{ minimizes in } \eta_{x,\lambda} U \cap \eta_{x,\lambda} N \text{ in case } S = \lim \eta_{x_i,\lambda_i\#} T$$

with $\lim x_i = x$ and $\lim \lambda_i = \lambda > 0$, and

(8) $$S \text{ minimizes in all of } \mathbb{R}^{n+\ell} \text{ in case } S = \lim \eta_{x_i,\lambda_i\#} T$$

with $\lim x_i = x$ and $\lim \lambda_i = 0$. (Cf. the discussion in the proof of 3.1(2).)

For convenience we define

(9) $$U_S = \begin{cases} \eta_{x,\lambda} U & \text{in case } \lim \lambda_i > 0 \text{ (as in (7))} \\ \mathbb{R}^{n+\ell} & \text{in case } \lim \lambda_i = 0 \text{ (as in (8))} \end{cases}$$

so that $S \in \mathcal{D}_n(U_S)$ for each $S \in \mathcal{T}$.

Now by definition one readily checks that

(10) $$\eta_{x,\lambda\#}\mathcal{T} = \mathcal{T}, \ 0 < \lambda < 1, \ |x| < 1 - \lambda,$$

and, by (4),

(11) $$\Theta^n(\mu_S, y) \le m - \beta \ \forall y \in U_S, \ S \in \mathcal{T}.$$

Furthermore by using the compactness theorem 2.4 together with the monotonicity 4.7 of Ch. 4, one readily checks that if $S_i \rightharpoonup S$ ($S_i, S \in \mathcal{T}$) and if $y, y_i \in B_1(0)$ with $\lim y_i = y$, then

(12) $$\Theta^n(\mu_S, y) \ge \limsup \Theta^n(\mu_{S_i}, y_i).$$

It now follows from (10), (11), (12) and 2.4 that all the hypotheses of Theorem A.4 (of Appendix A) are satisfied with

(13) $$\mathcal{F} = \{\varphi_S : S \in \mathcal{T}\} \quad \text{(using notation of Remark 3.4)}$$

and with sing defined by

(14) $$\operatorname{sing} \varphi_S = \{x \in U_S : \Theta^n(\mu_S, \cdot) \ge m - 1 + \beta\}$$

for $S \in \mathcal{T}$. We claim that in this case the additional hypothesis is satisfied with $d = n - 3$. Indeed suppose $d \ge n - 2$; then there is $S \in \mathcal{T}$ and $\eta_{y,\lambda\#} S = S \ \forall y \in L$, $\lambda > 0$ with L an $(n-2)$-dimensional subspace of $\mathbb{R}^{n+\ell}$, $L \subset \operatorname{sing} \varphi_S$. Since we can make a rotation of $\mathbb{R}^{n+\ell}$ to bring L into coincidence with $\mathbb{R}^{n-2} \times \{0\}$, we assume that $L = \mathbb{R}^{n-2} \times \{0\}$. Then by 3.5 we have

(15) $$S = [\![\mathbb{R}^{n-2}]\!] \times S_0,$$

where $S_0 \in \mathcal{D}_2(\mathbb{R}^N)$, $N = 2 + k$, with S_0 a 2-dimensional area minimizing cone in $\mathbb{R}^N$. Then $\operatorname{spt} S_0$ is contained in a finite union $\cup_{i=1}^q P_i$ of 2-planes, with $P_i \cap P_j = \{0\} \ \forall i \ne j$. (For a formal proof of this characterization of 2 dimensional area minimizing cones, see for example [Whi82].) In particular, since $\Theta^n(\mu_S, \cdot)$ is constant on $P_i \setminus \{0\}$ (by the Constancy 2.34 of Ch. 6), we have that $\Theta^n(\mu_S, y) \in \mathbb{Z}$ for *every* $y \in \mathbb{R}^{n+\ell}$, and by (11) it follows that $\Theta^n(\mu_S, y) \le m-1$ $\forall y \in \mathbb{R}^{n+\ell}$. That is, $\operatorname{sing}\varphi_S = \emptyset$, a contradiction, hence we can take $d = n-3$ as claimed. We have thus established (5) as required. □

Proof of 3.3(2): The proof goes similarly to 3.3(1). This time we assume (again without loss of generality) that

(1) $$U = B_2(0),$$

and we prove that T has a tangent plane at all points of $\operatorname{spt} T \cap B_1(0)$ except for a set $F \subset \operatorname{spt} T \cap B_1(0)$ with

(2) $$\mathcal{H}^{n-2+\alpha}(F) = 0 \ \forall \alpha > 0.$$

$\mathcal{T}$ is as described in the proof of 3.3(1), and for any $S \in \mathcal{T}$ and $\beta > 0$ we let

$$R_\beta(S) = \{x \in \operatorname{spt} S : B_\rho(x) \subset U_S \text{ and } h(\operatorname{spt} S, L, \rho, x) < \beta\rho \text{ for some } \rho > 0 \text{ and some } n\text{-dimensional subspace } L \text{ of } \mathbb{R}^{n+\ell}\},$$

where U_S is as in the proof of 3.3(1) (so that $S \in \mathcal{D}_n(U_S)$), and where we define

(3) $$h(\operatorname{spt} S, L, \rho, x) = \sup_{y \in \operatorname{spt} S \cap B_\rho(x)} |q(y-x)|,$$

with q the orthogonal projection of $\mathbb{R}^{n+\ell}$ onto $L^\perp$.

Now notice that (Cf. the proof of 3.3(1))

(4) $$\eta_{x,\lambda\#}\mathcal{T} = \mathcal{T} \ \forall\, 0 < \lambda < 1, \ |x| < 1-\lambda,$$

and

(5) $$\eta_{x,\lambda} R_\beta(S) = R_\beta(\eta_{x,\lambda\#}S), \ S \in \mathcal{T}.$$

Furthermore if $S_j \rightharpoonup S$, $S_j, S \in \mathcal{T}$, then by the monotonicity 4.7 of Ch. 4 it is quite easy to check that if $y \in R_\beta(S)$ and if $y_j \in \operatorname{spt} S_j$ with $y_j \to y$, then $y_j \in R_\beta(S_j)$ for all sufficiently large j. Because of this, and because of (4), (5) above, it is now straightforward to check that the hypotheses of A.4 hold with (again in notation of 3.4)

(6) $$\mathcal{F} = \{\varphi_S : S \in \mathcal{T}\}$$

and

(7) $$\operatorname{sing}\varphi_S = \operatorname{spt}\Theta^n(\mu_S, \cdot) \cap U_S \setminus R_\beta(S).$$

(Notice that $R_\beta(S)$ is completely determined by $\Theta^n(\mu_S, \cdot)$, and hence this makes sense.) In this case we claim that $d \le n-2$. Indeed if $d > n-2$ (i.e. $d = n-1$) then $\exists S \in \mathcal{T}$ such that

(8) $$\eta_{x,\lambda\#}S = S \ \forall x \in L, \ \lambda > 0, \text{ and } L \subset \operatorname{sing}\varphi_S$$

where L is an $(n-1)$-dimensional subspace. Then, supposing with loss of generality that $L = \mathbb{R}^{n-1} \times \{0\}$, we have by 3.5 that

(9) $$S = [\![\mathbb{R}^{n-1}]\!] \times S_0,$$

where S_0 is a 1-dimensional minimizing cone in $\mathbb{R}^{k+1}$. However it is easy to check that such a 1-dimensional minimizing cone necessarily has the form

(10) $$S_0 = m[\![\ell]\!],$$

where $m \in \mathbb{Z}$ and ℓ is a 1-dimensional subspace of $\mathbb{R}^{k+1}$. Thus (9) gives that $S = m[\![L]\!]$ where L is an n-dimensional subspace and hence $\operatorname{sing}\varphi_S = \emptyset$, a contradiction, so $d \le n-2$ as claimed.

We therefore conclude from A.4 that for each $S \in \mathcal{T}$

(11) $$\mathcal{H}^{n-2+\alpha}(\operatorname{spt} S \setminus R_\beta(S) \cap B_1(0)) = 0 \ \forall \alpha > 0.$$

If $\beta_j \downarrow 0$ we thus conclude in particular that

(12) $$\mathcal{H}^{n-2+\alpha}\big(\operatorname{spt} T \setminus \cup_{j=1}^\infty R_{\beta_j}(T) \cap B_1(0)\big) = 0 \ \forall \alpha > 0.$$

However by (1) we see that

(13) $$x \in \cup_{j=1}^\infty R_{\beta_j}(T) \iff T \text{ has a tangent plane at } x,$$

and therefore (12) gives (2) as required. □

4 Some Regularity Results (Arbitrary Codimension)

In this section, for $T \in \mathcal{D}_n(U)$ any integer multiplicity current, we define a relatively closed subset $\operatorname{sing} T$ of U by

4.1 $$\operatorname{sing} T = \operatorname{spt} T \setminus \operatorname{reg} T,$$

where $\operatorname{reg} T$ denotes the set of points $\xi \in \operatorname{spt} T$ such that for some $\rho > 0$ there is a $m \in \mathbb{Z} \setminus \{0\}$ and an embedded n-dimensional oriented C^1 submanifold M of $\mathbb{R}^{n+\ell}$ with $T = m[\![M]\!]$ in $B_\rho(\xi)$.

F.J. Almgren [Alm84] has proved the very important theorem that

$$\mathcal{H}^{n-2+\alpha}(\operatorname{sing} T) = 0 \; \forall \alpha > 0$$

in case $\operatorname{spt} T \subset N$, $\partial T = 0$ and T is minimizing in N, where N is a smooth embedded $(n+\ell)$-dimensional submanifold of $\mathbb{R}^{n+\ell}$. The proof is very non-trivial and requires development of a whole new range of results for minimizing currents. We here restrict ourselves to more elementary results.

Firstly, the following theorem is an immediate consequence of The Allard Theorem 5.2 of Ch. 5 and Lemma 1.2 of the present chapter.

4.2 Theorem. *Suppose $T \in \mathcal{D}_n(U)$ is integer multiplicity and minimizing in $U \cap N$ for some embedded C^2 $(n+\ell)$-dimensional submanifold N of $\mathbb{R}^{n+\ell}$, $(\overline{N} \setminus N) \cap U = \emptyset$, and suppose $\operatorname{spt} T \subset U \cap N$, $\partial T = 0$ (in U). Then $\operatorname{reg} T$ is dense in $\operatorname{spt} T$.*

(Note that by definition $\operatorname{reg} T$ is relatively open in $\operatorname{spt} T$.)

The following is a useful fact; however its applicability is limited by the hypothesis that $\Theta^n(\mu_T, y) = 1$.

4.3 Theorem. *Suppose $\{T_i\} \subset \mathcal{D}_n(U)$, $T \in \mathcal{D}_n(U)$ are integer multiplicity currents with T_i minimizing in $U \cap N_i$, T minimizing in $U \cap N$, N, N_i embedded $(n+\ell)$-dimensional C^2 submanifolds, and $\operatorname{spt} T_i \subset N_i$, $\operatorname{spt} T \subset N$, $\partial T_i = \partial T = 0$ (in U). Suppose also that N_i converges to N in the C^2 sense in U, $T_j \rightharpoonup T$ in $\mathcal{D}_n(U)$, and suppose $y \in N \cap U$ with $\Theta^n(\mu_T, y) = 1$, $y = \lim y_j$, where y_j is a sequence such that $y_j \in \operatorname{spt} T_j \; \forall j$. Then $y \in \operatorname{reg} T$ and $y_j \in \operatorname{reg} T_j$ for all sufficiently large j.*

Proof: By virtue of the monotonicity formula 4.7 of Ch. 4 (which is applicable by 1.2) it is easily checked that

$$\limsup \Theta^n(\mu_{T_j}, y_j) \le \Theta^n(\mu_T, y) = 1,$$

hence (since $\Theta^n(\mu_{T_j}, y_j) \ge 1$ by 4.8 of Ch. 4) we conclude that $\Theta^n(\mu_{T_j}, y_j) \to \Theta^n(\mu_T, y) = 1$. Hence by Allard's Theorem 5.2 of Ch. 5 we have $y \in \operatorname{reg} T$ and $y_j \in \operatorname{reg} T_j$ for all sufficiently large j. (1.2 justifies the use of 5.2 of Ch. 5.)

Next we have the following consequences of A.4 of Appendix A.

4.4 Theorem. *Suppose T is as in 4.2, and in addition suppose $\xi \in \operatorname{spt} T$ is such that $\Theta^n(\mu_T, \xi) < 2$. Then there is a $\rho > 0$ such that*

$$\mathcal{H}^{n-2+\alpha}(\operatorname{sing} T \cap B_\rho(\xi)) = 0 \; \forall \alpha > 0.$$

Proof: Let $\alpha = 2 - \Theta^n(\mu_T, \xi)$ and let $B_\rho(\xi)$ be such that $B_{2\rho}(\xi) \subset U$ and

(1) $$(\omega_n \sigma^n)^{-1} \mu_T(B_\sigma(\zeta)) < 2 - \alpha/2$$

$\forall \zeta \in \operatorname{spt} T \cap B_\rho(\xi)$, $0 < \sigma < \rho$. (Notice that such ρ exists by virtue of the monotonicity 4.7 of Ch. 4, which can be applied by Lemma 1.2.) Assume without loss of generality that $\xi = 0$, $\rho = 1$ and $U = B_2(0)$, and define $\mathcal{T}$ to be the set of weak limits S of sequences $\{S_i\}$ of the form $S_i = \eta_{x_i, \lambda_i \#} T$, $|x_i| < 1 - \lambda_i$, $0 < \lambda_i < 1$, where $\lim x_i$ and $\lim \lambda_i = \lambda$ are assumed to exist. Notice that

(2) $$\limsup \mathbb{M}_W(S_i) < \infty$$

for each $W \subset\subset \eta_{x,\lambda}(U)$ in case $\lambda > 0$ and for each $W \subset\subset \mathbb{R}^{n+\ell}$ in case $\lambda = 0$. Hence by the Compactness 2.4 any such S is integer multiplicity in U_S

(3) $$(U_S = \eta_{x,\lambda} U \text{ in case } \lambda > 0,\; U_S = \mathbb{R}^{n+\ell} \text{ in case } \lambda = 0)$$

and (Cf. the proof of 3.3(2))

(4) $$S \text{ minimizes in } \eta_{x,\lambda} U \cap \eta_{x,\lambda} N \text{ in case } \lambda > 0$$

(5) $$S \text{ minimizes in } \mathbb{R}^{n+\ell} \text{ in case } \lambda = 0.$$

One readily checks that, by definition of $\mathcal{T}$,

(6) $$\eta_{y,\tau\#} \mathcal{T} = \mathcal{T},\; 0 < \tau < 1,\; |y| < 1 - \tau$$

Furthermore we note that (by (1))

(7) $$\Theta^n(\mu_S, x) = 1,\; \mu_S\text{-a.e. } x \in U_S,$$

and by Allard's 5.2 of Ch. 5 there is $\delta > 0$ such that

(8) $$\operatorname{sing} S = \{x \in U_S : \Theta^n(\mu_S, x) \ge 1 + \delta\},\; S \in \mathcal{T}.$$

Now in view of (4), (5), (6), (7), (8) and the upper semi-continuity of Θ^n as in (12) in the proof of 3.3(1), all the hypotheses of A.4 of A are satisfied with $\mathcal{F} = \{\varphi_S : S \in \mathcal{T}\}$ (notation as in 3.4) and with $\operatorname{sing} \varphi_S = \{x \in U_S : \Theta^n(\mu_S, x) \ge 1 + \delta\}$ ($\equiv \operatorname{sing} S$ by (8)). In fact we claim that in this case we may take $d = n - 2$, because if $d = n - 1$ $\exists S \in \mathcal{T}$ and $\eta_{x,\lambda\#} S = S$ $\forall x \in L$, $\lambda > 0$, where $L \subset \operatorname{sing} S$ is an $(n-1)$-dimensional subspace of $\mathbb{R}^{n+\ell}$, then (Cf. the last part of the proof of 3.3 (2)) we have $S = m[\![Q]\!]$ for some n-dimensional subspace Q. Hence $\operatorname{sing} S = \emptyset$, a contradiction. □

The following theorem is often useful:

4.5 Theorem. *Suppose* $C \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$ *is minimizing in* $\mathbb{R}^{n+\ell}$, $\partial C = 0$, *and* C *is a cone:* $\eta_{0,\lambda\#}C = C\ \forall \lambda > 0$. *Suppose further that* $\operatorname{spt} C \subset \overline{H}$ *where* H *is an open* $\frac{1}{2}$*-space of* $\mathbb{R}^{n+\ell}$ *with* $0 \in \partial H$. *Then* $\operatorname{spt} C \subset \partial H$.

4.6 Remark: The reader will see that the theorem here is actually valid with any stationary rectifiable varifold V in $\mathbb{R}^{n+\ell}$ satisfying $\eta_{0,\lambda\#}V = V$ in place of C.

Proof of 4.5: Since the varifold V associated with C is stationary (by 1.2) in $\mathbb{R}^{n+\ell}$ we have 5.1 of Ch. 4 (Since $(Dr)^{\perp} = 0$ by virtue of the fact that C is a cone),

$$(1) \qquad \frac{d}{d\rho}\Big(\rho^{-n}\int_{\mathbb{R}^{n+\ell}} h\varphi(r/\rho)\,d\mu_C\Big) = \rho^{-n-1}\int_{\mathbb{R}^{n+\ell}} x\cdot(\nabla^C h)\varphi(r/\rho)\,d\mu_C$$

for each $\rho > 0$, where $r = |x|$ and φ is a non-negative C^1 function on $\mathbb{R}$ with compact support, and h is an arbitrary $C^1(\mathbb{R}^{n+\ell})$ function. ($\nabla^C h(x)$ denotes the orthogonal projection of $\nabla_{\mathbb{R}^{n+\ell}}h(x)$ onto the tangent space T_xV of V at x.)

Now suppose without loss of generality that $H = \{x = (x^1,\ldots,x^{n+\ell}) : x^1 > 0\}$ and select $h(x) \equiv x^1$. Then $x\cdot\nabla^C h = e_1^T\cdot x = e_1\cdot x^T = re_1\cdot\nabla^C r$, where v^T denotes orthogonal projection of v onto T_xV. Thus the term on the right side of (1) can be written $-\int_{\mathbb{R}^{n+\ell}}(e_1\cdot\nabla^C r)\big(r\varphi(r/\rho)\big)\,d\mu_C$, which in turn can be written $-\int_{\mathbb{R}^{n+\ell}} e_1\cdot\nabla^C\psi_\rho\,d\mu_C$, where $\psi_\rho(x) = \int_{|x|}^{\infty} r\varphi(r/\rho)\,dr$. (Thus ψ_ρ has compact support in $\mathbb{R}^{n+\ell}$.) But $e_1\cdot\nabla^C\psi_\rho \equiv \operatorname{div}_V(\psi_\rho e_1)$, and hence the term on the right of (1) actually vanishes by virtue of the fact that V is stationary. Thus (1) gives

$$\rho^{-n}\int_{\mathbb{R}^{n+\ell}} x_1\varphi(r/\rho)\,d\mu_C = \text{const.},\ 0 < \rho < \infty.$$

In view of the arbitrariness of φ, this implies

$$\rho^{-n}\int_{B_\rho(0)} x_1\,d\mu_C \equiv \text{const.}$$

However trivially we have $\lim_{\rho\downarrow 0}\rho^{-n}\int_{B_\rho(0)} x_1\,d\mu_C = 0$, and hence we deduce

$$\rho^{-n}\int_{B_\rho(0)} x_1\,d\mu_C = 0\ \forall\rho > 0.$$

Thus since $x_1 \ge 0$ on $\operatorname{spt} C$ ($\subset \overline{H}$), we conclude $\operatorname{spt} C \subset \partial H$ ($= \{x : x^1 = 0\}$). □

The following corollary of 4.5 follows directly by combining 4.5 and 3.1(2).

4.7 Corollary. *If* T *is as in 4.2, if* $\xi \in \operatorname{spt} T$, *if* Q *is a* C^1 *hypersurface in* $\mathbb{R}^{n+\ell}$ *such that* $\xi \in Q$ *and if* $\operatorname{spt} T$ *is locally on one side of* Q *near* ξ, *then all tangent cones* C *of* T *at* ξ *satisfy* $\operatorname{spt} C \subset T_\xi Q \cap T_\xi N$.

5 Codimension 1 Theory

We begin by looking at those integer multiplicity currents $T \in \mathcal{D}_n(U)$ with $\operatorname{spt} T \subset N \cap U$, N an $(n+1)$-dimensional oriented embedded submanifold of $\mathbb{R}^{n+\ell}$ with $(\overline{N}\setminus N)\cap U = \emptyset$ and such that

$$\textbf{5.1} \qquad T = \partial[\![E]\!]$$

(in U), where E is an $\mathcal{H}^{n+1}$-measurable subset of N. (We know by 3.16 of Ch. 6, 1.4 that all minimizing currents $T \in \mathcal{D}_n(U)$ with $\partial T = 0$ and $\operatorname{spt} T$ in N can be locally decomposed into minimizing currents of this special form.)

5.2 Remark: The fact that T has the form 5.1 and T is integer multiplicity evidently is equivalent to the requirement that if $V \subset U$ is open, and if φ is a C^2 diffeomorphism of V onto an open subset of $\mathbb{R}^{n+\ell}$ such that $\varphi(V\cap N) = G$, G open in $\mathbb{R}^{n+1}$, then $\varphi(E)$ has locally finite perimeter in G. This is an easy consequence of 2.35 of Ch. 6, and in fact we see from this and 4.3 of Ch. 3 that any T of the form 5.1 with $\mathbb{M}_W(T) < \infty\ \forall W \subset\subset U$ is automatically integer multiplicity with

$$(\ddagger) \qquad \Theta^n(T,x) = 1,\ \mu_T\text{-a.e. } x \in U.$$

We shall here develop the theory of minimizing currents of the form 5.1; indeed we show this is naturally done using only the more elementary facts about currents. In particular we shall not in this section have any need of the Compactness 3.11 of Ch. 6 (instead we use only the elementary BV Compactness Theorem 2.6 of Ch. 2), nor shall we need the Deformation Theorem and the subsequent material of Chapter 6.

The following theorem could be derived from the general Compactness 2.4, but here (as we mentioned above) we can give a more elementary treatment. In this theorem, and subsequently, we take $U \subset \mathbb{R}^{n+\ell}$ to be open and $\mathcal{O}$ will denote the collection of $(n+1)$-dimensional oriented embedded C^2 submanifolds N of $\mathbb{R}^{n+\ell}$ with $(\overline{N}\setminus N)\cap U = \emptyset$, $N\cap U \ne \emptyset$. A sequence $\{N_j\} \subset \mathcal{O}$ is said to converge to $N \in \mathcal{O}$ in the C^2 sense in U if there are orientation preserving C^2 embeddings $\psi_j : N\cap U \to N_j$ with $\psi_j \to 1_{N\cap U}$ then $\eta_{x,\lambda}N$ converges to T_xN in the C^2 sense in W as $\lambda \downarrow 0$, for each $W \subset\subset \mathbb{R}^{n+\ell}$.

In the following theorem p is a proper C^2 map $U \to N\cap U$ such that in some neighborhood $V \subset U$ of $N\cap U$, p coincides with the nearest point projection of V onto N. (Since the nearest point projection is C^2 in some neighborhood of $N\cap U$ it is clear that such p exists.)

5.3 (Compactness Theorem for minimizing T as in 5.1). *Suppose $T_j \in \mathcal{D}_n(U)$, $T_j = \partial[\![E_j]\!]$ (in U), E_j $\mathcal{H}^{n+1}$-measurable subsets of $N_j \cap U$, $N_j \in \mathcal{O}$, $N_j \to N \in \mathcal{O}$ in the C^2 sense described above, and suppose T_j is integer multiplicity and minimizing in $U \cap N_j$.*

Then there is a subsequence $\{T_{j'}\}$ with $T_{j'} \rightharpoonup T$ in $\mathcal{D}_n(U)$, T integer multiplicity, $T = \partial[\![E]\!]$ (in U), $\chi_{p(E_{j'})} \to \chi_E$ in $L^1_{\text{loc}}(\mathcal{H}^{n+1}, U)$, $\mu_{T_{j'}} \to \mu_T$ (in the usual sense of Radon measures) in U, and T is minimizing in $N \cap U$.

5.4 Remarks: (1) Recall (from 5.2) that the hypothesis that T_j is integer multiplicity is automatic if we assume merely that $\mathbb{M}_W(T_j) < \infty \; \forall W \subset\subset U$.

(2) We make no *a-priori* assumptions on local boundedness of the mass of T_j (we see in the proof that this is automatic for minimizing currents as in 5.1).

(3) Let $h(x,t) = x + t(p(x) - x)$, $x \in U$, $0 \le t \le 1$. Using the homotopy formula 2.25 of Ch. 6 (and in particular the inequality 2.27 of Ch. 6) together with the fact that $N_j \to N$ in the C^2 sense in U, it is straightforward to check that

$$T_j - T = \partial R_j, \; R_j = h_\#([\![(0,1)]\!] \times T_j) + p_\#[\![E_j]\!] - [\![E]\!]$$

with

$$\mathbb{M}_W(R_{j'}) \to 0 \;\forall W \subset\subset U,$$

provided that $\chi_{p(E_{j'})} \to \chi_E$ as claimed in the theorem. Thus once we establish $\chi_{p(E_{j'})} \to \chi_E$ for some E, then we can use the argument of 2.4 (with $S_j = 0$) in order to conclude

(i) T is minimizing in U

(ii) $\mu_{T_{j'}} \to \mu_T$ in U.

(Notice we have not had to use the deformation theorem here.) In the following proof we therefore concentrate on proving $\chi_{p(E_{j'})} \to \chi_E$ in $L^1_{\text{loc}}(\mathcal{H}^{n+1}, N \cap U)$ for some subsequence $\{j'\}$ and some E such that $\partial[\![E]\!]$ has locally finite mass in U. (T is then automatically integer multiplicity by 5.2.)

Proof of 5.3: We first establish a local mass bound for the T_j in U: if $\xi \in N$ and $B_{\rho_0}(\xi) \subset U$, then

$$\text{(1)} \qquad \mathbb{M}(T_j \llcorner B_\rho(\xi)) \le \tfrac{1}{2}\mathcal{H}^n(\partial B_\rho(\xi) \cap N), \; \mathcal{L}^1\text{-a.e. } \rho \in (0, \rho_0).$$

This is proved by simple area comparison as follows:

With $r(x) = |x - \xi|$, by the elementary slicing theory of 4.5(1),(2) of Ch. 6 we have that, for $\mathcal{L}^1$-a.e. $\rho \in (0, \rho_0)$, the slice $\langle [\![E_j]\!], r, \rho\rangle$ (i.e. the slice of $[\![E_j]\!]$ by $\partial B_\rho(\xi)$) is integer multiplicity, and (using $T_j = \partial[\![E_j]\!]$),

$$\partial[\![E_j \cap B_\rho(\xi)]\!] = T_j \llcorner B_\rho(\xi) + \langle [\![E_j]\!], r, \rho\rangle.$$

Hence (applying ∂ to this identity)

$$\partial(T_j \llcorner B_\rho(\xi)) = -\partial\langle [\![E_j]\!], r, \rho\rangle, \; \mathcal{L}^1\text{-a.e. } \rho \in (0, \rho_0),$$

and by Definition 1.1 of minimizing

$$\mathbb{M}(T_j \llcorner B_\rho(\xi)) \le \mathbb{M}\langle [\![E_j]\!], r, \rho\rangle.$$

Since $-\widetilde{T}_j$ is also minimizing in $N \cap U$ we then have

$$\text{(2)} \qquad \mathbb{M}(T_j \llcorner B_\rho(\xi)) \le \min\left\{\mathbb{M}\langle [\![E_j]\!], r, \rho\rangle, \langle [\![\widetilde{E}_j]\!], r, \rho\rangle\right\}$$

for $\mathcal{L}^1$-a.e. $\rho \in (0, \rho_0)$, where $\widetilde{E}_j = N \cap U \setminus E_j$.

Now of course $[\![\widetilde{E}]\!] + [\![E_j]\!] = [\![N \cap U]\!]$, so that (for a.e. $\rho \in (0, \rho_0)$)

$$\langle [\![E_j]\!], r, \rho\rangle + \langle [\![\widetilde{E}_j]\!], r, \rho\rangle = \langle N, r, \rho\rangle$$

and hence (2) gives (1) as required (because $\mathbb{M}(\langle N, r, \rho\rangle) \le \mathcal{H}^n(N \cap \partial B_\rho(\xi))$ by virtue of the fact that $|Dr| = 1$, hence $|\nabla^N r| \le 1$).

Now by virtue of (1) and 5.2 we deduce from the BV Compactness 2.6 of Ch. 2 that some subsequence $\{\chi_{p(E_{j'})}\}$ of $\{\chi_{p(E_j)}\}$ converges in $L^1_{\text{loc}}(\mathcal{H}^{n+1}, N \cap U)$ to χ_E, where $E \subset N$ is $\mathcal{H}^{n+1}$-measurable and such that $\partial[\![E]\!]$ is integer multiplicity (in U). The remainder of the theorem now follows as described in 5.4(3). □

5.5 (Existence of tangent cones). *Suppose $T = \partial[\![E]\!] \in \mathcal{D}_n(U)$ is integer multiplicity, with $E \subset N \cap U$, $N \in \mathcal{O}$, and T is minimizing in $U \cap N$. Then for each $x \in \operatorname{spt} T$ and each sequence $\{\lambda_j\} \downarrow 0$ there is a subsequence $\{\lambda_{j'}\}$ and an integer multiplicity $C \in \mathcal{D}_n(\mathbb{R}^{n+\ell})$ with C minimizing in $\mathbb{R}^{n+\ell}$, $0 \in \operatorname{spt} C \subset T_xN$, $\Theta^n(\mu_C, 0) = \Theta^n(\mu_T, x)$, $C = \partial[\![F]\!]$, F an $\mathcal{H}^{n+1}$-measurable subset of T_xN,*

$$\text{(1)} \qquad \mu_{\eta_{x,\lambda_{j'}\#}T} \to \mu_C \text{ in } \mathbb{R}^{n+\ell}, \; \chi_{p\left(\eta_{x,\lambda_j}(E)\right)} \to \chi_F \text{ in } L^1_{\text{loc}}(\mathcal{H}^{n+1}, T_xN),$$

where p is the orthogonal projection of $\mathbb{R}^{n+\ell}$ onto T_xN, and

$$\text{(2)} \qquad \eta_{0,\lambda\#}C = C, \; \eta_{0,\lambda}F = F \;\forall \lambda > 0.$$

5.6 Remark: The proof given here is independent of the general tangent cone Existence 3.1.

Proof of 5.5: As we remarked prior to 5.3, $\eta_{x,\lambda_j}N$ converges to T_xN in the C^2 sense in W for each $W \subset\subset \mathbb{R}^{n+\ell}$. By the Compactness 5.3 we then have a subsequence λ_j, such that all the required conclusions, except possibly for 5.5(2) and the fact that $0 \in \operatorname{spt} C$, hold. To check that $0 \in \operatorname{spt} C$ and that 5.5(2) is valid, we first note by 1.2 that the varifold V associated with T is stationary in $N \cap U$ (and that V therefore has locally bounded generalized mean curvature $\underline{H}$ in $N \cap U$). Therefore by the monotonicity formula 4.7 of Ch. 4, and by 4.8 of Ch. 4, we have

$$\Theta^n(\mu_V, x) \text{ exists and is } \geq 1.$$

Since $\mu_{\eta_{x,\lambda_j\#}T} \to \mu_C$, we then have $\Theta^n(\mu_C, 0) = \Theta^n(\mu_T, x) \geq 1$, so $0 \in \operatorname{spt} C$, and by Theorem 6.1 of Ch. 4 we deduce that the varifold V_C associated with C is a cone. Then in particular $x \wedge \vec{C}(x) = 0$ for μ_C-a.e. $x \in \mathbb{R}^{n+\ell}$ and hence, if we let h be the homotopy $h(t,x) = tx + (1-t)\lambda x$, we have $h_\#([\![(0,1)]\!] \times C) = 0$, and then by the homotopy formula 2.25 of Ch. 6 (since $\partial C = 0$) we have $\eta_{0,\lambda\#}C = C$ as required. Finally since $\operatorname{spt} C$ has locally finite $\mathcal{H}^n$-measure (indeed by 4.8 of Ch. 4 $\operatorname{spt} C$ is the closed set $\{y \in \mathbb{R}^{n+\ell} : \Theta^n(\mu_C, y) \geq 1\}$), we have

$$[\![F]\!] = [\![\widetilde{F}]\!],$$

where $\widetilde{F}$ is the (open) set $\{y \in T_xN \setminus \operatorname{spt} C : \Theta^{n+1}(\mathcal{H}^{n+1}, T_xN, y) = 1\}$. Evidently $\eta_{0,\lambda}(\widetilde{F}) = \widetilde{F}$ (because $\eta_{0,\lambda}(\operatorname{spt} C) = \operatorname{spt} C$). Hence the required result is established with $\widetilde{F}$ in place of F. □

5.7 Corollary.[4] *Suppose T is as in* 5.5 *and in addition suppose there is an n-dimensional submanifold Σ embedded in $\mathbb{R}^{n+\ell}$ with $x \in \Sigma \subset N \cap U$ for some $x \in \operatorname{spt} T$, and suppose* $\operatorname{spt} T \setminus \Sigma$ *lies locally, near x, on one side of Σ. Then $x \in \operatorname{reg} T$.* ($\operatorname{reg} T$ *is as in 4.1)*

Proof:l Let $C = \partial[\![F]\!]$ ($F \subset T_xN$) be any tangent cone for T at x. By assumption $\operatorname{spt}[\![F]\!] \subset \overline{H}$, where H is an open $\frac{1}{2}$-space in T_xN with $0 \in \partial H$. Then, by 4.5, $\operatorname{spt} C \subset \partial H$ and hence the Constancy 2.34 of Ch. 6 since C is integer multiplicity rectifiable, it follows that $C = \pm\partial[\![H]\!]$. However $\operatorname{spt}[\![F]\!] \subset \overline{H}$, hence $C = +\partial[\![H]\!]$. Then $\Theta^n(\mu_C, y) \equiv 1$ for $y \in \partial H$, and in particular $\Theta^n(\mu_C, 0)$ $(= \Theta^n(\mu_T, x)) = 1$, so that $x \in \operatorname{reg} T$ (by Allard's Theorem 5.2 of Ch. 5) as required.

We next want to prove the main regularity theorem for codimension 1 currents. We continue to define $\operatorname{sing} T$, $\operatorname{reg} T$ as in 4.1.

5.8 Theorem. *Suppose $T = \partial[\![E]\!] \in \mathcal{D}_n(U)$ is integer multiplicity, with $E \subset N \cap U$, $n \in \mathcal{O}$, and T minimizing in $N \cap U$. Then $\operatorname{sing} T = \emptyset$ for $n \leq 6$, $\operatorname{sing} T$ is locally finite in U for $n = 7$, and $\mathcal{H}^{n-7+\alpha}(\operatorname{sing} T) = 0$ $\forall \alpha > 0$ in case $n > 7$.*

Proof: We are going to use the abstract dimension reducing argument of Appendix A (Cf. the proof of 4.4).

[4] Cf. Miranda [Mir67]

To begin we note that it is enough (by re-scaling, translation, and restriction) to assume that

$$U = \breve{B}_2(0)$$

and to prove that

$$\begin{cases} \operatorname{sing} T \cap B_1(0) = \emptyset \text{ if } n \leq 6, \ \operatorname{sing} T \cap B_1(0) \text{ discrete if } n = 7, \\ \mathcal{H}^{n-7+\alpha}(\operatorname{sing} T \cap B_1(0)) = 0 \ \forall \alpha > 0 \text{ if } n > 7. \end{cases}$$

Let $\mathcal{T}$ be the set of currents as defined in the proof of 4.4[5], and for each $S \in \mathcal{T}$ let φ_S be the function : $\mathbb{R}^{n+\ell} \to \mathbb{R}^{n+1}$ associated with S as in 3.4. Also, let

$$\mathcal{F} = \{\varphi_S : S \in \mathcal{T}\}$$

and define

$$\operatorname{sing} \varphi_S = \operatorname{sing} S.$$

($\operatorname{sing} S$ as defined in 4.1.)
By A.4 we then have either $\operatorname{sing} S = \emptyset$ for all $S \in \mathcal{T}$ (and hence $\operatorname{sing} T = \emptyset$) or

$$\dim B_1(0) \cap \operatorname{sing} S \leq d,$$

where $d \in [0, n-1]$ is the integer such that

$$\dim B_1(0) \cap \operatorname{sing} S \leq d \text{ for all } S \in \mathcal{T}$$

and such that there is $S \in \mathcal{T}$ and a d-dimensional subspace L of $\mathbb{R}^{n+\ell}$ such that

$$\eta_{x,\lambda\#}S = S \ \forall x \in L, \ \lambda > 0$$

and

5.9 $$\operatorname{sing} S = L.$$

Supposing without loss of generality that $L = \mathbb{R}^d \times \{0\}$, we then (by 3.5) have

5.10 $$S = [\![R^d]\!] \times S_0$$

where $\partial S_0 = 0$, S_0 is minimizing in $\mathbb{R}^{n+\ell-1}$, and $\operatorname{sing} S_0 = \{0\}$. (With S as in 5.10, $\operatorname{sing} S_0 = \{0\} \iff$ 5.9.) Also, by definition of $\mathcal{T}$, $\operatorname{spt} S \subset$ some $(n+1)$-dimensional subspace of $\mathbb{R}^{n+\ell}$, hence without loss of generality we have that S_0 is an $(n-d)$-dimensional minimizing cone in $\mathbb{R}^{n-d+1}$ with $\operatorname{sing} S_0 = \{0\}$. Then by

[5] We still have $\Theta^n(\mu_S, x) - 1$ for μ_S-a.e. $x \in U_S$, this time by 5.3 and 5.2 (‡)

the result of J. Simons (see B) we have $n - d > 6$; i.e. $d \le n - 7$. Notice that this contradicts $d \ge 0$ in case $n < 7$. Thus for $n < 7$ we must have $\operatorname{sing} T = \emptyset$ as required. If $n = 7$, $\operatorname{sing} T$ is discrete by the last part of A.4.

5.11 Corollary. *If T is as in 5.8, and if $T_1 \in \mathcal{D}_n(U)$ is obtained by equipping a component of $\operatorname{reg} T$ with multiplicity 1 and with orientation of T, then $\partial T_1 = 0$ (in U) and T_1 is minimizing in $U \cap N$.*

5.12 Remark: Notice that this means we can write

$$T = \sum_{j=1}^{\infty} T_j,$$

where each T_j is obtained by equipping a component M_j of $\operatorname{reg} T$ with multiplicity 1 and with the orientation of T; then $M_i \cap M_j = \emptyset \ \forall i \neq j$, $\partial T_j = 0$, and T_j is minimizing in $U \ \forall j$. Furthermore (since $\mu_{T_j}(B_\rho(x)) \ge c\rho^n$ for $B_\rho(x) \subset U$ and $x \in \operatorname{spt} T_j$ by virtue of 1.2 and the monotonicity 4.7 of Ch. 4) only finitely many T_j can have support intersecting a given compact subset of U.

Proof of 5.11: The main point is to prove

(1) $$\partial T_1 = 0 \text{ in } U.$$

The fact that T_1 is minimizing in U will then follow from 1.4 and the fact that $\mathbb{M}_W(T_1) + \mathbb{M}_W(T - T_1) = \mathbb{M}_W(T) \ \forall W \subset\subset U$.

To check (1) let $\omega \in \mathcal{D}^{n-1}(U)$ be arbitrary and note that if $\zeta \equiv 0$ in some neighborhood of $\operatorname{spt} T \setminus M_1$

(2) $$T_1(d(\zeta\omega)) = T(d(\zeta\omega)) = \partial T(\zeta\omega) = 0.$$

Now corresponding to any $\varepsilon > 0$ we construct ζ as follows: since $\mathcal{H}^{n-1}(\operatorname{sing} T) = 0$ (by 5.8) and since $\operatorname{sing} T \cap \operatorname{spt} \omega$ is compact, we can find a finite collection of balls $\{B_{\rho_j}(\xi_j)\}_{j=1,\dots,P}$ with $\xi_j \in \operatorname{sing} T \cap \operatorname{spt} \omega$ and $\sum_{j=1}^{P} \rho_j^{n-1} < \varepsilon$. For each $j = 1, \dots, P$ let $\varphi_j \in C_c^\infty(\mathbb{R}^{n+\ell})$ be such that $\varphi_j \equiv 1$ on $B_{\rho_j}(\xi_j)$, $\varphi_j = 0$ on $\mathbb{R}^{n+\ell} \setminus B_{2\rho_j}(\xi_j)$, and $0 \le \varphi_j \le 1$ everywhere. Now choose $\zeta = \prod_{j=1}^{P} \varphi_j$ in a neighborhood of $\operatorname{spt} T_1$ and so that $\zeta \equiv 0$ in a neighborhood of $\operatorname{spt} T \setminus \operatorname{spt} T_1$. Then $d\zeta = \sum_{i=1}^{P} \prod_{j \neq i}^{P} \varphi_j \, d\varphi_i$ on $\operatorname{spt} T_1$, and hence

$$|d(\zeta\omega) - \zeta d\omega| \le c|\omega| \sum_{j=1}^{P} \rho_j^{n-1} \le c\varepsilon|\omega| \text{ on } \operatorname{spt} T_1.$$

The letting $\varepsilon \downarrow 0$ in (2), and noting that $\zeta d\omega \to d\omega$ $\mathcal{H}^n$-a.e. in $\operatorname{spt} T_1 \cap N \cap \operatorname{spt} \omega$ (and using $|\zeta| \le 1$), we conclude $T_1(d\omega) = 0$. That is $\partial T_1 = 0$ in U as required. □

Finally we have the following lemma.

5.13 Lemma. *If $T_1 = \partial[\![E_1]\!]$, $T_2 = \partial[\![E_2]\!] \in \mathcal{D}_n(U)$, U bounded, $E_1, E_2 \subset U \cap N$, N of class C^4, $N \in \mathcal{O}$, T_1, T_2 minimizing in $U \cap N$, $\operatorname{reg} T_1$, $\operatorname{reg} T_2$ are connected, and $E_1 \cap V \subset E_2 \cap V$ for some neighborhood V of ∂U, then $\operatorname{spt}[\![E_1]\!] \subset \operatorname{spt}[\![E_2]\!]$ and either $[\![E_1]\!] = [\![E_2]\!]$ or $\operatorname{spt} T_1 \cap \operatorname{spt} T_2 \subset \operatorname{sing} T_1 \cap \operatorname{sing} T_2$.*

Proof: Since $\mathcal{H}^{n+1}(\operatorname{spt} T_j) = 0$ (in fact $\operatorname{spt} T_j$ has locally finite $\mathcal{H}^n$-measure in U by virtue of the fact that $\Theta^n(\mu_{T_j}, x) \ge 1 \ \forall x \in \operatorname{spt} T_j$), we may assume that E_1 and E_2 are open with $U \cap \partial E_j = U \cap \partial \overline{E}_j = \operatorname{spt} T_j$, $j = 1, 2$.

Let $S_1, S_2 \in \mathcal{D}_n(U)$ be the currents defined by

$$S_1 = \partial[\![E_1 \cap E_2]\!], \ S_2 = \partial[\![E_1 \cup E_2]\!].$$

Using the hypothesis concerning V we have

(3) $$S_j \llcorner (V \cap U) = T_j \llcorner (V \cap U), \ j = 1, 2.$$

On the other hand we trivially have

$$[\![E_1 \cap E_2]\!] + [\![E_1 \cup E_2]\!] = [\![E_1]\!] + [\![E_2]\!],$$

so (applying ∂) we get

(4) $$S_1 + S_2 = T_1 + T_2.$$

Furthermore $E_1 \cap E_2 \subset E_1 \cup E_2$, so

(5) $$\begin{aligned} \mathbb{M}_W(S_1) + \mathbb{M}(S_2) &= \mathbb{M}_W(S_1 + S_2) \\ &= \mathbb{M}_W(T_1 + T_2) \quad \text{(by (4))} \\ &\le \mathbb{M}_W(T_1) + \mathbb{M}_W(T_2) \end{aligned}$$

$\forall W \subset\subset U$. On the other hand, choosing an open V_0 so that $\partial U \subset V_0 \subset\subset V$, and using (3) together with the fact that T_1 is minimizing, we have

$$\mathbb{M}_W(S_1) \ge \mathbb{M}_W(T_1), \ W = U \setminus \overline{V}_0,$$

and hence (combining this with (5))

$$\mathbb{M}_W(S_2) \le \mathbb{M}_W(T_2)$$

for $W = U \setminus \overline{V}_0$. Thus (using (3) with $j = 2$) S_2 is minimizing in U. Likewise S_1 is minimizing in U.

We next want to prove that either $T_1 = T_2$ or $\operatorname{reg} T_1 \cap \operatorname{reg} T_2 = \emptyset$. Suppose $\operatorname{reg} T_1 \cap \operatorname{reg} T_2 \neq \emptyset$. If the tangent spaces of $\operatorname{reg} T_1$ and $\operatorname{reg} T_2$ coincide at every point of their intersection, then using suitable local coordinates $(x, z) \in \mathbb{R}^n \times \mathbb{R}$ for N near a point $\xi \in \operatorname{reg} T_1 \cap \operatorname{reg} T_2$, we can write

$$\operatorname{reg} T_j = \operatorname{graph} u_j, \ j = 1, 2,$$

where $Du_1 = Du_2$ at each point where $u_1 = u_2$, and where both u_1, u_2 are (weak) C^1 solutions of the equation

$$\frac{\partial}{\partial x_i}\Big(\frac{\partial F}{\partial p_i}(x, u, Du)\Big) - \frac{\partial F}{\partial z}(x, u, Du) = 0,$$

where $F = F(x, z, p)$, $(x, z, p) \in \mathbb{R}^n \times \mathbb{R} \times \mathbb{R}^n$, is the area functional for graphs $z = u(x)$ relative to the local coordinates x, z for N. Since N is C^4 we then deduce (from standard quasilinear elliptic theory—see e.g. [GT01]) that u_1, u_2 are $C^{3,\alpha}$. Now the difference $u_1 - u_2$ of the solutions evidently satisfies an equation of the general form

$$D_j(a_{ij} D_i u) + b_i D_i u + cu = 0,$$

where a_{ij}, b_i, c are $C^{2,\alpha}$. By standard unique continuation results (see e.g. [Pro60]) we then see that $Du_1 = Du_2$ at each point where $u_1 = u_2$ is impossible if $u_1 - u_2$ changes sign. On the other hand the Harnack inequality ([GT01]) tells us that either $u_1 \equiv u_2$ or $|u_1 - u_2| > 0$ in case $u_1 - u_2$ does not change sign. Thus we deduce that either $T_1 = T_2$ or $\operatorname{reg} T_1 \cap \operatorname{reg} T_2 = \emptyset$ or there is a point $\xi \in \operatorname{reg} T_1 \cap \operatorname{reg} T_2$ such that $\operatorname{reg} T_1$ and $\operatorname{reg} T_2$ intersect *transversely* at ξ. But then we would have $\mathcal{H}^{n-1}(\operatorname{sing} \partial[\![E_1 \cap E_2]\!]) > 0$, which by virtue of 5.8 contradicts the fact (established above) that $\partial[\![E_1 \cap E_2]\!]$ is minimizing in U.

Thus either $T_1 = T_2$ or $\operatorname{reg} T_1 \cap \operatorname{reg} T_2 = \emptyset$, and it follows in either case that $E_1 \subset E_2$. On the other hand we then have $\operatorname{sing} T_1 \cap \operatorname{reg} T_2 = \emptyset$ and $\operatorname{sing} T_2 \cap \operatorname{reg} T_1 = \emptyset$ by virtue of 5.7. Thus we conclude that $E_1 \subset E_2$ and $\operatorname{spt} T_1 \cap \operatorname{spt} T_2 \subset \operatorname{sing} T_1 \cap \operatorname{sing} T_2$ as required.

Chapter 8

Theory of General Varifolds

1 Basics, First Rectifiability Theorem

We let $G(n+\ell, n)$ denote the collection of all n-dimensional subspaces of $\mathbb{R}^{n+\ell}$, equipped with the metric $\rho(S, T) = |p_S - p_T| = \big(\sum_{i,j=1}^{n+\ell}(p_S^{ij} - p_T^{ij})^2\big)^{\frac{1}{2}}$ where p_S, p_T denote the orthogonal projections of $\mathbb{R}^{n+\ell}$ onto S, T respectively, and $p_S^{ij} = e_i \cdot p_S(e_j)$, $p_T^{ij} = e_i \cdot p_T(e_j)$ are the corresponding matrices with respect to the standard orthonormal basis $e_1, \ldots, e_{n+\ell}$ for $\mathbb{R}^{n+\ell}$.

For a subset $A \subset \mathbb{R}^{n+\ell}$ we define

$$G_n(A) = A \times G(n+\ell, n),$$

equipped with the product metric. Of course then $G_n(K)$ is compact for each compact $K \subset \mathbb{R}^{n+\ell}$. $G_n(\mathbb{R}^{n+\ell})$ is locally homeomorphic to a Euclidean space of dimension $n + \ell + nk$.

By an n-varifold we mean simply any Radon measure V on $G_n(\mathbb{R}^{n+\ell})$. By an n-varifold on U (U open in $\mathbb{R}^{n+\ell}$) we mean any Radon measure V on $G_n(U)$. Given such an n-varifold V on U, there corresponds a Radon measure $\mu = \mu_V$ on U (called the *weight* of V) defined by

$$\mu(A) = V(\pi^{-1}(A)), \ A \subset U,$$

where, here and subsequently, π is the projection $(x,S)\mapsto x$ of $G_n(U)$ onto U. The *mass* $\mathbb{M}(V)$ of V is defined by

$$\mathbb{M}(V)=\mu_V(U)\quad (=V(G_n(U))).$$

for any Borel subset $A\subset U$ we use the usual terminology $V\llcorner G_n(A)$ to denote the restriction of V to $G_n(A)$; thus

$$(V\llcorner G_n(A))(B)=V(B\cap G_n(A)),\ B\subset G_n(U).$$

Given an n-rectifiable varifold $\underline{v}(M,\theta)$ on U (in the sense of Ch. 4) there is a corresponding n-varifold V (also denoted by $\underline{v}(M,\theta)$, or simply $\underline{v}(M)$ in case $\theta\equiv 1$ on M), defined by

$$V(A)=\mu(\pi(TM\cap A)),\ A\subset G_n(U),$$

where $\mu=\mathcal{H}^n\llcorner\theta$ and $TM=\{(x,T_xM):x\in M_*\}$, with M_* the set of $x\in M$ such that M has an approximate tangent space T_xM with respect to θ at x in the sense of 1.7 of Ch. 3. Evidently V, so defined, has weight measure $\mu_V=\mathcal{H}^n\llcorner\theta=\mu$.

The question of when a general n-varifold actually corresponds to an n-rectifiable varifold in this way is satisfactorily answered in the next theorem. Before stating this we need a definition:

1.1 Definition: Given $T\in G(n+\ell,n)$, $x\in U$, and $\theta\in(0,\infty)$, we say that an n-varifold V on U has tangent space T with multiplicity θ at x if

$$(\ddagger)\qquad \lim_{\lambda\downarrow 0}V_{x,\lambda}=\theta\underline{v}(T),$$

where the limit is in the usual sense of Radon measures on $G_n(\mathbb{R}^{n+\ell})$. In 1.1 ($\ddagger$) we use the notation that $V_{x,\lambda}$ is the n-varifold defined by

$$V_{x,\lambda}(A)=\lambda^{-n}V\big(\{(\lambda y+x,S):(y,S)\in A\}\cap G_n(U)\big)$$

for $A\subset G_n(\mathbb{R}^{n+\ell})$.

1.2 (First Rectifiability Theorem.) *Suppose V is an n-varifold on U which has a tangent space T_x with multiplicity $\theta(x)\in(0,\infty)$ for μ_V-a.e. $x\in U$. Then V is n-rectifiable; in fact $M\equiv\{x\in\operatorname{spt}V: T_x$ and $\theta(x)$ both exist$\}$ is $\mathcal{H}^n$-measurable, countably n-rectifiable, θ is locally $\mathcal{H}^n$-integrable on M, and $V=\underline{v}(M,\theta)$.*

In the proof of 1.2 (and also subsequently) we shall need the following technical lemma:

1.3 Lemma. *Let V be any n-varifold on U. Then for μ_V-a.e. $x\in U$ there is a Radon measure $\eta_V^{(x)}$ on $G(n+\ell,n)$ such that, for any continuous β on $G(n+\ell,n)$,*

$$\int_{G(n+\ell,n)}\beta(S)\,d\eta_V^{(x)}(S)=\lim_{\rho\downarrow 0}\frac{\int_{G_n(B_\rho(x))}\beta(S)\,dV(y,S)}{\mu_V(B_\rho(x))}.$$

Furthermore for any Borel set $A\subset U$,

$$\int_{G_n(A)}\beta(S)\,dV(x,S)=\int_A\int_{G(n+\ell,n)}\beta(S)\,d\eta_V^{(x)}(S)d\mu_V(x)$$

provided $\beta\ge 0$.

Proof: The proof is a simple consequence of the differentiation theory for Radon measures and the separability of $\mathcal{K}(X,\mathbb{R})$ (notation as in §4 of Ch. 1) for compact separable metric spaces X. Specifically, write $\mathcal{K}=\mathcal{K}(G(n+\ell,n),\mathbb{R})$, $\mathcal{K}^+=\{\beta\in\mathcal{K}:\beta\ge 0\}$, and let $\beta_1,\beta_2,\ldots\in\mathcal{K}^+$ be dense in $\mathcal{K}^+$. By the XXX Theorem 3.23 of Ch. 1 we know that (since there is a Radon measure γ_j on $\mathbb{R}^{n+\ell}$ characterized by $\gamma_j(B)=\int_{G_n(B)}\beta_j(S)\,dV(y,S)$ for Borel sets $B\subset\mathbb{R}^{n+\ell}$)

$$(1)\qquad e(x,j)=\lim_{\rho\downarrow 0}\frac{\int_{G_n(B_\rho(x))}\beta_j(S)\,dV(y,S)}{\mu_V(B_\rho(x))}$$

exists for each $x\in\mathbb{R}^{n+\ell}\setminus Z_j$, where Z_j is a Borel set with $\mu_V(Z_j)=0$ and $e(x,j)$ is a μ_V-measurable function of x, with

$$(2)\qquad \int_A e(x,j)\,d\mu_V(x)=\int_{G_n(A)}\beta_j(S)\,dV(y,S)$$

for any Borel set $A\subset\mathbb{R}^{n+\ell}$.

Now let $\varepsilon>0$, $\beta\in\mathcal{K}^+$, $x\in\mathbb{R}^{n+\ell}\setminus(\cup_{j=1}^\infty Z_j)$, and choose β_j such that $\sup|\beta-\beta_j|<\varepsilon$. Then for any $\rho>0$

(3)
$$\Bigg|\frac{\int_{G_n(B_\rho(x))}\beta(S)\,dV(y,S)}{\mu_V(B_\rho(x))}-\frac{\int_{G_n(B_\rho(x))}\beta_j(S)\,dV(y,S)}{\mu_V(B_\rho(x))}\Bigg|\le\varepsilon\frac{V(G_n(B_\rho(x)))}{\mu_V(B_\rho(x))}=\varepsilon,$$

and hence by (1) we conclude that

$$(4)\qquad \widetilde{\eta}_V^{(x)}(\beta)\equiv\lim_{\rho\downarrow}\frac{\int_{G_n(B_\rho(x))}\beta(S)\,dV(y,S)}{\mu_V(B_\rho(x))}$$

exists for all $\beta\in\mathcal{K}^+$ and all $x\in\mathbb{R}^{n+\ell}\setminus(\cup_{j=1}^\infty Z_j)$. Of course, since $|\widetilde{\eta}_V^{(x)}(\beta)|\le\sup|\beta|\ \forall\beta\in\mathcal{K}^+$, by the Riesz Representation Theorem 4.14 of Ch. 1 we have

$\widetilde{\eta}_V^{(x)}(\beta) = \int_{G(n+\ell,n)} \beta(S)\, d\eta_V^{(x)}(S)$, where $\eta_V^{(x)}$ is the total variation measure associated with $\widetilde{\eta}_V^{(x)}$.

Finally the last part of the lemma follows directly from (2), (3) if we keep in mind that $e(x,j)$ in (1) is exactly $\widetilde{\eta}_V^{(x)}(\beta_j) \int_{G(n+\ell,n)} \beta_j(S)\, d\eta_V^{(x)}(S)$ □

We are now able to give the proof of 1.2.

Proof of 1.2: By definition 1.1, μ_V has approximate tangent space T_x with multiplicity $\theta(x)$ in the sense of 1.7 of Ch. 3 for μ_V-a.e. $x \in U$. Hence by 1.9 of Ch. 3 we have that M is $\mathcal{H}^n$-measurable countably n-rectifiable, θ is locally $\mathcal{H}^n$-integrable on M and in fact $\mu_V = \mathcal{H}^n \llcorner \theta$ in U (if we set $\theta \equiv 0$ in $U \setminus M$).

Now if $x \in M$ is one of the μ_V-almost all points such that $\eta_V^{(x)}$ exists, and if β is a non-negative continuous function on $G(n+\ell,n)$, then we evidently have $\eta_V^{(x)}(\beta) = \theta(x)\beta(T_x)$ and hence by the second part of 1.3 we have

(1) $$\int_{G_n(A)} \beta(S)\, dV(x,S) = \int_{M\cap A} \beta(T_x)\, d\mu_V(x)$$

for any Borel set $A \subset U$. From the arbitrariness of A and β it then easily follows that

(2) $$\int_{G_n(U)} f(x,S)\, dV(x,S) = \int_M f(x,T_x)\, d\mu_V(x)$$

for any non-negative $f \in C_c(G_n(U))$, and hence we have shown $V = \underline{v}(M,\theta)$ as required (because $\mu_V = \mathcal{H}^n \llcorner \theta$ as mentioned above). □

2 First Variation

We can make sense of first variation for a general varifold V on U. We first need to discuss *mapping* of such a general n-varifold. Suppose $U, \widetilde{U}$ open $\subset \mathbb{R}^{n+\ell}$ and $f : U \to \widetilde{U}$ is C^1 with $f|\operatorname{spt}\mu_V \cap U$ *proper*. Then we define the image varifold $f_\# V$ on $\widetilde{U}$ by

2.1 $$f_\# V(A) = \int_{F^{-1}(A)} J_S f(x)\, dV(x,S), \ A \text{ Borel}, \ A \subset G_n(\widetilde{U}),$$

where $F : G_n^+(U) \to G_n(\widetilde{U})$ is defined by $F(x,S) = (f(x), df_x(S))$ and where

$$J_S f(x) = \left(\det\left((df_x|S)^* \circ (df_x|S)\right)\right)^{\frac{1}{2}}, \ (x,S) \in G_n(U).$$
$$G_n^+(U) = \{(x,S) \in G_n(U) : J_S f(x) \neq 0\}.$$

(Notice that this agrees with our previous definition given in §1 of Ch. 4 in case $V = \underline{v}(M,\theta)$.)

Now given any n-varifold V on U we define the *first variation* δV of V, which is a linear functional on $\mathcal{K}(U,\mathbb{R}^{n+\ell})$ (notation as in §4 of Ch. 1) by

2.2 $$\delta V(X) = \frac{d}{dt}\mathbb{M}\left(\varphi_{t\#} V \llcorner G_n(K)\right)\Big|_{t=0},$$

where $\{\varphi_t\}_{-1<t<1}$ is any 1-parameter family as in 5.6 of Ch. 2 (and K compact is as in 5.6 of Ch. 2). Of course we can compute $\delta V(X)$ explicitly by differentiation under the integral in 2.1. This gives (by *exactly* the computations in §5 of Ch. 2)

2.3 $$\delta V(X) = \int_{G_n(U)} \operatorname{div}_S X(x)\, dV(x,S),$$

where, for any $S \in G(n+\ell,n)$,

2.4 $$\operatorname{div}_S X = \textstyle\sum_{i=1}^{n+\ell} \nabla_i^S x^i = \sum_{i=1}^n \langle \tau_i, D_{\tau_i} X\rangle,$$

where $\tau_1,\ldots,\tau_n$ is an orthonormal basis for S and $\nabla_i^S = e_i \cdot \nabla^S$, with $\nabla^S f(x) = p_S(\nabla_{\mathbb{R}^{n+\ell}} f(x))$, $f \in C^1(U)$. (p_S is the orthogonal projection of $\mathbb{R}^{n+\ell}$ onto S.)

By analogy with 2.4 of Ch. 4 we then say that V is *stationary in* U if $\delta V(X) = 0\ \forall X \in \mathcal{K}(U,\mathbb{R}^{n+\ell})$.

More generally V is said to have *locally bounded first variation in* U if for each $W \subset\subset U$ there is a constant $c < \infty$ such that $|\delta V(X)| \le c \sup_U |X|\ \forall X \in \mathcal{K}(U,\mathbb{R}^{n+\ell})$ with $\operatorname{spt}|X| \subset W$. Evidently, by the general Riesz Representation 4.14 of Ch. 1, this is equivalent to the requirement that there is a Radon measure $\|\delta\|$ (the total variation measure of δV) on U characterized by

2.5 $$\|\delta V\|(W) = \sup_{X\in\mathcal{K}(U,\mathbb{R}^{n+\ell}),\, |X|\le 1,\, \operatorname{spt}|X|\subset W} |\delta V(X)| \quad (<\infty)$$

for any open $W \subset\subset U$. Notice that then by 4.14 of Ch. 1 we can write

2.6 $$\delta V(X) = \int_{G_n(U)} \operatorname{div}_S X(x)\, dV(x,S) \equiv -\int_U \nu\cdot X\, d\|\delta V\|,$$

where ν is $\|\delta V\|$-measurable with $|\nu| = 1$ $\|\delta V\|$-a.e. in U. By XXX Theorem 3.23 of Ch. 1 we know furthermore that

2.7 $$D_{\mu_V}\|\delta V\|(x) \equiv \lim_{\rho\downarrow 0} \frac{\|\delta V\|(B_\rho(x))}{\mu_V(B_\rho(x))}$$

exists μ_V-a.e. and that (writing $\underline{H}(x) = D_{\mu_V}\|\delta V\|(x)\nu(x)$)

2.8 $$\int_U \nu\cdot X\, d\|\delta V\| = \int_U \underline{H}\cdot X\, d\mu_V + \int_U \nu\cdot X\, d\sigma,$$

with

2.9 $\sigma = \|\delta V\| \llcorner Z,\ Z = \{x \in U : D_{\mu_V}\|\delta V\|(x) = +\infty\}.\ (\mu_V(Z) = 0.)$

Thus we can write

2.10
$$\begin{aligned}\delta V(x) &= \int_{G_n(U)} \operatorname{div}_S X(x)\, dV(x,S)\\ &= -\int_U \underline{H}\cdot X\, d\mu_V - \int_Z \nu\cdot X\, d\sigma\end{aligned}$$

for $X \in \mathcal{K}(U,\mathbb{R}^{n+\ell})$.

By analogy with the classical identity 4.30 of Ch. 2 we call $\underline{H}$ *the generalized mean curvature of* V, Z *the generalized boundary of* V, σ *the generalized boundary measure of* V, and $\nu|Z$ *the generalized unit co-normal of* V.

3 Monotonicity and Consequences

In this section we assume that V is an n-varifold in U with locally bounded first variation in U (as in 2.5).

We first consider a point $x \in U$ such that there is $0 < \rho_0 < \operatorname{dist}(x,\partial U)$ and $\Lambda \ge 0$ with

3.1 $$\|\delta V\|(B_\rho(x)) \le \Lambda\mu_V(B_\rho(x)),\ 0 < \rho < \rho_0.$$

Subject to 3.1 we can choose (in 2.3 of Ch. 4) $X_y = \gamma(r)(y-x)$, $r = |y-x|$, $y \in U$ as in §4 of Ch. 4 and note that (by essentially the same computation as in §4 of Ch. 4)

3.2 $$\operatorname{div}_S X = n\gamma(r) + r\gamma'(r)\textstyle\sum_{i,j=1}^{n+\ell} e_S^{ij}\dfrac{x^i-y^i}{r}\dfrac{x^j-y^j}{r},$$

where (e_S^{ij}) is the matrix of the orthogonal projection p_S of $\mathbb{R}^{n+\ell}$ onto the n-dimensional subspace S. We can then take $\gamma(r) = \varphi(r/\rho)$ (again as in §4 of Ch. 4) and, noting that $\sum_{i,j=1}^{n+\ell} e_S^{ij}\frac{x^i-y^i}{r}\cdot\frac{x^j-y^j}{r} = 1 - \left|p_{S^\perp}\left(\frac{y-x}{r}\right)\right|^2$, conclude (Cf. 4.7 of Ch. 4 with $\alpha = 1$) that $e^{\Lambda\rho}\rho^{-n}\mu_V(B_\rho(x))$ is increasing in ρ, $0 < \rho < \rho_0$, and, for $0 < \sigma \le \rho < \rho_0$,

3.3
$$\begin{aligned}\Theta^n(\mu_v,x) &\le e^{\Lambda\sigma}\omega_n^{-1}\sigma^{-n}\mu_V(B_\sigma(x)) \le e^{\Lambda\rho}\omega_n^{-1}\rho^{-n}\mu_V(B_\rho(x))\\ &\quad - \omega_n^{-1}\int_{G_n(B_\rho(x)\setminus B_\sigma(x))} r^{-n-2}|p_{S^\perp}(y-x)|^2\, dV(y,S).\end{aligned}$$

In fact if $\Lambda = 0$ (so that V is stationary in $B_{\rho_0}(x)$) we get the precise identity

3.4 $$\Theta^n(\mu_v,x) = \omega_n^{-1}\rho^{-n}\mu_V(B_\rho(x)) - \omega_n^{-1}\int_{G_n(B_\rho(x))} r^{-n-2}|p_{S^\perp}(y-x)|^2\, dV(y,S)$$

for $0 < \rho < \rho_0$.

Using $X_y = h(y)\gamma(r)(y-x)$ $(r = |y-x|)$ in 2.3 of Ch. 4 we also deduce that the following analogue of 5.1 of Ch. 4:

3.5
$$\begin{aligned}\frac{d}{d\rho}\left(\rho^{-n}\widetilde{I}(\rho)\right) &= \rho^{-n}\frac{d}{d\rho}\int |p_{S^\perp}(y-x)/r|^2\varphi(r/\rho)h(y)\, dV(y,S)\\ &\quad + \rho^{-n-1}\left(\delta V(X) + \int (y-x)\cdot\nabla^S h(y)\varphi(r/\rho)\, dV(y,S)\right),\end{aligned}$$

where $\widetilde{I}(\rho) = \int \varphi(r/\rho)h\, d\mu_V$.

3.6 Lemma. *Suppose V has locally bounded first variation in U. Then for μ_V-a.e. $x \in U$, $\Theta^n(\mu_V,x)$ exists and is real-valued; in fact $\Theta^n(\mu_V,x)$ exists whenever there is a constant $\Lambda(x) < \infty$ such that*

$$\|\delta V\|(B_\rho(x)) \le \Lambda(x)\mu_V(B_\rho(x)),\ 0 < \rho < \frac{1}{2}\operatorname{dist}(x,\partial U).$$

(Such a constant $\Lambda(x)$ exists for μ_V-a.e. $x \in U$ by virtue of XXX Theorem 3.23 of Ch. 1.)

Furthermore $\Theta^n(\mu_V,x)$ is a μ_V-measurable function of x.

Proof: The first part of the lemma follows directly from the monotonicity formula 3.3. The μ_V-measurability of $\Theta^n(\mu_V,\cdot)$ follows from the fact that $\mu_V(B_\rho(x)) \ge \limsup_{y\to x}\mu_V(B_\rho(y))$, which guarantees that $\mu_V(B_\rho(x))/(\omega_n\rho^n)$ is Borel measurable and hence μ_V-measurable for each fixed ρ. Since

$$\Theta^n(\mu_V,x) = \lim_{\rho\downarrow 0}(\omega_n\rho^n)^{-1}\mu_V(B_\rho(x))$$

for μ_V-a.e. $x \in U$, we then have μ_V-measurability of $\Theta^n(\mu_V,\cdot)$ as claimed.

3.7 Theorem. (Semi-continuity of Θ^n under varifold convergence.) *Suppose $V_i \to V$ (as Radon measures in $G_n(U)$) and $\Theta^n(V_i,y) \ge 1$ except on a set $B_i \subset U$ with $\mu_{V_i}(B_i \cap W) \to 0$ for each $W \subset\subset U$, and suppose that each V_i has locally bounded first variation in U with $\liminf\|\delta V_i\|(W) < \infty$ for each $W \subset\subset U$. Then $\|\delta V\|(W) \le \liminf\|\delta V_i\|(W)$ $\forall W \subset\subset U$ and $\Theta^n(\mu_V,y) \ge 1$ μ_V-a.e. in U.*

3.8 Remarks: (1) The fact that $\|\delta V\|(W) \le \liminf\|\delta V_i\|(W)$ is a trivial consequence of the definitions of $\|\delta V\|$, $\|\delta V_i\|$ and the fact that $V_i \to V$, so we have only to prove the last conclusion that $\Theta^n(\mu_V,y) \ge 1$ μ_V-a.e.

(2) The proof that $\Theta^n(\mu_V, y) \ge 1$ μ_V-a.e. to be given below is slightly complicated; the reader should note that if $\|\delta V\| \le \Lambda \mu_V$ in U (i.e. if V has generalized boundary measure $\sigma = 0$ and bounded $\underline{H}$—see 2.10 above—then the result is a very easy consequence of the monotonicity formula 3.3.

Proof of 3.7: Set $\mu_i = \mu_{V_i}$, $\mu = \mu_V$, and take any $W \subset\subset U$ and $\rho_0 \in (0, \operatorname{dist}(W, \partial U))$. For $i, j \ge 1$, consider the set $A_{i,j}$ consisting of all points $y \in W \setminus B_i$ such that

$$(1) \qquad \|\delta V_i\|(B_\rho(y)) \le j\mu_i(B_\rho(y)),\ 0 < \rho < \rho_0,$$

and let $B_{i,j} = W \setminus A_{i,j}$. Then if $x \in B_{i,j}$ we have *either* $x \in B_i \cap W$ or

$$(2) \qquad \mu_i(B_\sigma(x)) \le j^{-1}\|\delta V_i\|(B_\sigma(x)) \text{ for some } \sigma \in (0, \rho_0).$$

Let $\mathcal{B}$ be the collection of balls $B_\sigma(x)$ with $x \in B_{i,j}$, $\sigma \in (0, \rho_0)$, and with (2) holding. By the Besicovitch Covering Lemma (§3.12 of Ch. 1) there are families $\mathcal{B}_1, \ldots, \mathcal{B}_N \subset \mathcal{B}$ with $N = N(n+\ell)$, with $B_{i,j} \setminus B_i \subset \cup_{\ell=1}^N (\cup_{B \in \mathcal{B}_\ell} B)$ and with each $\mathcal{B}_\ell$ a pairwise disjoint family. Hence if we sum in (2) over balls $B \in \cup_{\ell=1}^N \mathcal{B}_\ell$, we get

$$\mu_i(B_{i,j}) \le N j^{-1}\|\delta V_i\|(\widetilde{W}) + \mu_i(B_i \cap W)$$

($\widetilde{W} = \{x \in U : \operatorname{dist}(x, W) < \rho_0\}$), so

$$\mu_i(B_{i,j}) \le cj^{-1} + \mu_i(B_i \cap W),$$

with c independent of i, j. In particular for each $i, j \ge 1$

$$(3) \qquad \mu(\operatorname{interior}(\cap_{\ell=i}^\infty B_{\ell,j})) \le \liminf_{q\to\infty} \mu_q(\operatorname{interior}(\cap_{\ell=i}^\infty B_{\ell,j})) \le cj^{-1},$$

since $\mu_q(B_q \cap W) \to 0$ as $q \to \infty$.

Now let $j \in \{1, 2, \ldots\}$ and consider the possibility that there is a point $x \in W$ such that $x \in W \setminus \operatorname{interior}(\cap_{q=i}^\infty B_{q,j})$ for *each* $i = 1, 2, \ldots$. Then we could select, for each $i = 1, 2, \ldots$, $y_i \in W \setminus \cap_{q=i}^\infty B_{q,j}$ with $|y_i - x| < 1/i$. Thus there are sequences $y_i \to x$ and $q_i \to \infty$ such that $y_i \notin B_{q_i,j}$ for each $i = 1, 2, \ldots$. Then $y_i \in A_{q_i,j}$ and hence (by (1))

$$\|\delta V_{q_i}\|(B_\rho(y_i)) \le j\mu_{q_i}(B_\rho(y_i)),\ 0 < \rho < \rho_0,$$

for all $i = 1, 2, \ldots$. Then by the monotonicity formula 3.3 (with $\Lambda = j$) together with the fact that $\Theta^n(\mu_{q_i}, y_i) \ge 1$ we have

$$\mu_{q_i}(B_\rho(y_i)) \ge e^{-j\rho}\omega_n\rho^n,\ 0 < \rho < \rho_0,$$

so that $\Theta^n(\mu, x) \ge 1$ for such an x. Thus we have proved $\Theta^n(\mu, x) \ge 1$ for each x with $x \in W \setminus (\cup_{i=1}^\infty \operatorname{interior}(\cap_{\ell=i}^\infty B_{\ell,j}))$ for some $j \in \{1, 2, \ldots\}$. That is

$$(4) \qquad \Theta^n(\mu, x) \ge 1\ \forall x \in W \setminus (\cap_{j=1}^\infty \cup_{i=1}^\infty \operatorname{interior}(\cap_{\ell=i}^\infty B_{\ell,j})).$$

However

$$(5) \qquad \begin{aligned} \mu(\cap_{j=1}^\infty \cup_{i=1}^\infty \operatorname{interior}(\cap_{\ell=i}^\infty B_{\ell,j})) &\le \mu(\cup_{i=1}^\infty \operatorname{interior}(\cap_{\ell=i}^\infty B_{\ell,j}))\ \forall j \ge 1 \\ &= \lim_{i\to\infty} \mu(\operatorname{interior}(\cap_{\ell=i}^\infty B_{\ell,j})) \\ &\le cj^{-1} \text{ by (3)}, \end{aligned}$$

so $\mu(\cap_{j=1}^\infty \cup_{i=1}^\infty \operatorname{interior}(\cap_{\ell=i}^\infty B_{\ell,j})) = 0$ and the theorem is established (by (4)). □

4 Constancy Theorem

4.1 (Constancy Theorem.) *Suppose V is an n-varifold in U, V is stationary in U, and $U \cap \operatorname{spt} \mu_V \subset M$, where M is a connected n-dimensional C^2 submanifold of $\mathbb{R}^{n+\ell}$. Then $V = \theta_0 \underline{v}(M)$ for some constant θ_0.*

4.2 Remarks: (1) Notice in particular this implies $(\overline{M} \setminus M) \cap U = \emptyset$ (if $V \ne 0$); this is not *a-priori* obvious from the assumptions of the theorem.

(2) J. Duggan in his PhD thesis [Dug86] has extended 4.1 to the case when M is merely Lipschitz.

(3) The reader will see that, with only minor modifications to the proof to be given below, the theorem continues to hold if N is an embedded $(n+\ell)$-dimensional C^2 submanifold of $\mathbb{R}^{n+\ell}$ and if V is stationary in $U \cap N$ in the sense that $\delta V(X) = 0\ \forall X \in \mathcal{K}(U; \mathbb{R}^{n+\ell})$ with $X_x \in T_x N\ \forall x \in N$, *provided we are given* $\operatorname{spt} V \subset \{(x, S) : x \in N \text{ and } S \subset T_x N\}$. (This last is equivalent to $\operatorname{spt} \mu_V \subset N$ and $p_\# V = V$, where $p : U \to U \cap N$ coincides with the nearest point projection onto $U \cap N$ in some neighborhood of $U \cap N$.)

Proof of 4.1: We first want to argue that $V = \underline{v}(M, \theta)$ for some positive locally $\mathcal{H}^n$-integrable function θ on M.

To do this first take any $f \in C_c^2(U)$ with $M \subset \{x \in U : f(x) = 0\}$ and note that by 2.3

$$(1) \qquad \delta V(f\nabla f) = \int |p_S(\nabla f)|^2\, dV(x, S),$$

because (using notation as in 2.3)

$$\begin{aligned}\operatorname{div}_S(f\nabla f) &= \nabla^S f\cdot\nabla f + f\operatorname{div}_S\nabla f\\ &= |p_S(\nabla f)|^2 \text{ on } M,\end{aligned}$$

where we used $f\equiv 0$ on M. Since $\delta V=0$, we conclude from (1) that

(2) $$p_S(\nabla f(x))=0 \quad\text{for all}\quad (x,S)\in\operatorname{spt}V.$$

Now let $\xi\in M$ be arbitrary. We can find an open $W\subset U$ with $\xi\in W$ and such that there are $C^2_c(U)$ functions $f_1,\ldots,f_k$ with $M\subset\cap_{j=1}^k\{x: f_j(x)=0\}$ and with $(T_xM)^\perp$ being exactly the space spanned by $\nabla f_1(x),\ldots,\nabla f_k(x)$ for each $x\in M\cap W$. (One easily checks that such W and $f_1,\ldots,f_k$ exists.) Then (2) implies that

(3) $$p_S((T_xM)^\perp)=0 \quad\text{for all}\quad (x,S)\in G_n(W)\cap\operatorname{spt}V.$$

But (3) says exactly that $S=T_xM$ for all $(x,S)\in G_n(W)\cap\operatorname{spt}V$, so that (since ξ was an arbitrary point of M), we have

(4) $$\int f(x,S)\,dV(x,S)=\int_{M\cap U} f(x,T_xM)\,d\mu_(x),\ f\in C_c(G_n(U)),$$

On the other hand we know from monotonicity 3.3 that $\theta(x)\equiv\Theta^n(\mu_V,x)$ exists for all $x\in M\cap U$, and hence (since $\Theta^n(\mathcal{H}^n\llcorner M,x)=1$ for each $x\in M$, by smoothness of M), we can use the XXX Theorem 3.23 of Ch. 1 to conclude from (4) that in fact

(5) $$\int f(x,S)\,dV(x,S)=\int_{M\cap U} f(x,T_xM)\theta(x)\,d\mathcal{H}^n(x),\ f\in C_c(G_n(U)),$$

(so that $V=\underline{v}(M,\theta)$ as required).

It thus remains only to prove that $\theta=$ const. on $M\cap U$. Since M is C^2 we can take $X\in\mathcal{K}(U,\mathbb{R}^{n+\ell})$ such that $X_x\in T_xM\ \forall x\in M\cap U$. Then by (5) and 2.3 $\delta V(X)=0$ is just the statement that $\int_{M\cap U}\operatorname{div}X\,\theta\,d\mathcal{H}^n=0$, where $\operatorname{div}X$ is the classical divergence of $X|M$ in the usual sense of differential geometry. Using local coordinates (in some neighborhood $\widetilde U\subset\mathbb{R}^n$) this tells us that

$$\int_{\widetilde U}\sum_{i=1}^n\frac{\partial X_i}{\partial x_i}\widetilde\theta\,d\mathcal{L}^n=0 \quad\text{if}\quad X_i\in C^1_c(\widetilde U),\ i=1,\ldots,n,$$

where $\widetilde\theta$ is θ expressed in terms of the local coordinates. In particular

$$\int_{\widetilde U}\frac{\partial\zeta}{\partial x_i}\widetilde\theta\,d\mathcal{L}^n=0\ \forall\zeta\in C_c(U),\ i=1,\ldots,n$$

and it is then standard that $\widetilde\theta=$ constant in $\widetilde U$. Hence (since M is connected) θ is constant in M. □

5 Varifold Tangents and Rectifiability Theorem

Let V be a n-varifold in U and let x be any point of U such that

5.1 $$\Theta^n(\mu_V,x)=\theta_0\in(0,\infty)\quad\text{and}\quad\lim_{\rho\downarrow 0}\rho^{1-n}\|\delta V\|(B_\rho(x))=0.$$

By definition of δV (in 2) and the Compactness Theorem 4.16 of Ch. 1 for Radon measures, we can select a sequence $\lambda_j\downarrow 0$ such that $\eta_{x,\lambda_j\#}V$ converges (in the sense of Radon measures) to a varifold C such that

$$C \text{ is stationary in } \mathbb{R}^{n+\ell}$$

and

5.2 $$\frac{\mu_C(B_\rho(x))}{\omega_n\rho^n}\equiv\theta_0\ \forall\rho>0.$$

Since $\delta C=0$ we can use 5.2 together with the monotonicity formula 3.4 to conclude

$$\int_{G_n(B_\rho(0))}\frac{|p_{S^\perp}(x)|^2}{|x|^{n+2}}\,dC(x,S)=0\ \forall\rho>0,$$

so that $p_{S^\perp}(x)=0$ for C-a.e. $(x,S)\in G_n(\mathbb{R}^{n+\ell})$, and hence $p_{S^\perp}(x)=0$ for *all* $(x,S)\in\operatorname{spt}C$ by continuity of $p_{S^\perp}(x)$ in (x,S). Then by the same argument as in the proof of Theorem 6.1 of Ch. 4, except that we use 3.5 in place of 5.1 of Ch. 4, we deduce that μ_C satisfies

5.3 $$\lambda^{-n}\mu_C(\eta_{0,\lambda}(A))=\mu_C(A),\ A\subset\mathbb{R}^{n+\ell},\ \lambda>0.$$

We would *like* to prove the stronger result $\eta_{0,\lambda\#}C=C$ (which of course implies 5.3, but we are only able to do this in case $\Theta^n(\mu_C,x)>0$ for μ_C-a.e. x (see 5.7 below). Whether of not $\eta_{0,\lambda\#}C=C$ without the additional hypothesis on $\Theta^n(\mu_C,\cdot)$ seems to be an open question.

5.4 Definition: Given V and x as in 5.1 we let $\operatorname{Var}\operatorname{Tan}(V,x)$ ("the varifold tangent of V at x") be the collection of all $C=\lim\eta_{x,\lambda_j\#}V$ obtained as described above.

Notice that by the above discussion any $C\in\operatorname{Var}\operatorname{Tan}(V,x)$ is stationary in $\mathbb{R}^{n+\ell}$ and satisfies 5.3.

The following rectifiability theorem for n-varifolds is a central part of the theory of n-varifolds with locally bounded first variation.

5.5 Theorem (Rectifiability Theorem.) *Suppose V has locally bounded first variation in U and $\Theta^n(\mu_V,x)>0$ for μ_V-a.e. $x\in U$. Then V is an n-rectifiable varifold.*

(Thus $V = \underline{v}(M,\theta)$, with M a $\mathcal{H}^n$-measurable countably n-rectifiable subset of U and θ a non-negative locally $\mathcal{H}^n$-integrable function on U.)

5.6 Remark: We are going to use 1.2. In fact we show that V has a tangent plane (in the sense of 1.1) at the point x where (i) $\Theta^n(\mu_V, x) > 0$, (ii) $\eta_V^{(x)}$ (as in 1.3) exists, (iii) $\Theta^n(\mu_V, \cdot)$ is μ_V-approximately continuous at x, and (iv) $\|\delta V\|(B_\rho(x)) \le \Lambda(x)\mu_V(B_\rho(x))$ for $0 < \rho < \rho_0 = \min\{1, \operatorname{dist}(x, \partial U)\}$. Since conditions (i)–(iv) all hold μ_V-a.e. in U (notice that (iii) holds μ_V-a.e. by virtue of the μ_V-measurability of $\Theta^n(\mu_V, \cdot)$ proved in 3.6), the required rectifiability of V will then follow from 1.2

Before beginning the proof of 5.5 we give the following important corollary.

5.7 Corollary. *Suppose $x \in U$, 5.1 holds, and $\lim_{\rho\downarrow 0} \rho^{-n}\mu_V(\{y \in B_\rho(x) : \Theta^n(\mu_V, y) < 1\}) = 0$. If $C \in \operatorname{Var}\operatorname{Tan}(V,x)$, then C is rectifiable and*

$$\eta_{0,\lambda\#}C = C \;\forall \lambda > 0.$$

Proof: From the hypothesis $\rho^{-n}\mu_V(\{y \in B_\rho(x) : \Theta^n(\mu_V, y) < 1\}) \to 0$ and the Semi-continuity 3.7, we have $\Theta^n(\mu_C, y) \ge 1$ for μ_C-a.e. $y \in \mathbb{R}^{n+\ell}$. Hence by 5.5 we have that C is n-rectifiable. On the other hand, since $\Theta^n(\mu_C, y) = \Theta^n(\mu_C, \lambda y)\ \forall \lambda > 0$ (by 5.3), we can write $C = \underline{v}(M,\theta)$ with $\eta_{0,\lambda}(M) = M\ \forall\lambda > 0$ and $\theta(\lambda y) = \theta(y)\ \forall \lambda > 0$, $y \in \mathbb{R}^{n+\ell}$. (Viz. simply set $\theta(y) = \Theta^n(\mu_C, y)$ and $M = \{y \in \mathbb{R}^{n+\ell} : \theta(y) > 0\}$.) It then trivially follows that $y \in T_y M$ whenever the approximate tangent space $T_y M$ exists, and hence $\eta_{0,\lambda\#}C = C$ as required. □

Proof of 5.5: Let x be as in 5.6(i)-(iv) and take $C \in \operatorname{Var}\operatorname{Tan}(V,x)$. (We know $\operatorname{Var}\operatorname{Tan}(V,x) \ne \emptyset$ because 5.6(i), (iv) imply 5.1.) Then C is stationary in $\mathbb{R}^{n+\ell}$ and

(1) $$\frac{\mu_c(B_\rho(0))}{\omega_n\rho^n} \equiv \theta_0 \;\forall \rho > 0 \quad (\theta_0 = \Theta^n(\mu_V, x)).$$

Also for any $y \in \mathbb{R}^{n+\ell}$ (using (1) and the monotonicity formula 3.3)

$$\begin{aligned}\frac{\mu_C(B_\rho(y))}{\omega_n\rho^n} &\le \frac{\mu_C(B_R(y))}{\omega_n R^n} \le \frac{\mu_C(B_{R+|y|}(0))}{\omega_n(R+|y|)^n}(1+|y|/R)^n \\ &= \theta_0(1+|y|/R)^n \to \theta_0 \text{ as } R\uparrow\infty.\end{aligned}$$

That is (again using the monotonicity formula 3.3),

(2) $$\Theta^n(\mu_C, y) \le \frac{\mu_C(B_\rho(y))}{\omega_n\rho^n} \le \theta_0 \;\forall y \in \mathbb{R}^{n+\ell},\ \rho > 0.$$

Now let $V_j = \eta_{x,\lambda_j\#}V$, where $\lambda_j \downarrow 0$ is such that $\lim \eta_{x,\lambda_j\#}V = C$ and where we are still assuming x is as in 5.6(i)–(iv).

From 5.6(iii) we have (with $\varepsilon(\rho)\downarrow 0$ as $\rho\downarrow 0$)

(3) $$\Theta^n(\mu_V, y) \ge \theta_0 - \varepsilon(\rho),\ y \in G\cap B_\rho(x),$$

where $G \subset U$ is such that

(4) $$\mu_V(B_\rho(x)\setminus G) \le \varepsilon(\rho)\rho^n,\ \rho \text{ sufficiently small.}$$

Taking $\rho = \lambda_j$ we see that (3), (4) imply

(5) $$\Theta^n(\mu_{V_j}, y) \le \theta_0 - \varepsilon_j,\ y \in G_j \cap B_1(0)$$

with G_j such that

(6) $$\mu_{V_j}(B_1(0)\setminus G_j) \le \varepsilon_j,$$

where $\varepsilon_j \to 0$ as $j\to\infty$. Thus, using (5), (6) and the semi-continuity result of 3.7, we obtain

(7) $$\Theta^n(\mu_C, y) \ge \theta_0 \text{ for } \mu_C\text{-a.e. } y \in \mathbb{R}^{n+\ell}$$

(and hence for *every* $y \in \operatorname{spt}\mu_C$ by 3.4). Then by combining (2) and (7) we have

$$\Theta^n(\mu_C, y) \equiv \theta_0 \equiv \frac{\mu_C(B_\rho(y))}{\omega_n\rho^n}\;\forall y \in \operatorname{spt}\mu_C,\ \rho > 0.$$

Then by the monotonicity formula 3.4 (with $V = C$), we have

$$p_{S^\perp}(x-y) = 0 \text{ for } C\text{-a.e. } (x,S) \in G_n(\mathbb{R}^{n+\ell}).$$

Thus (using the continuity of $p_{S^\perp}(x-y)$ in (x,S)) we have

(8) $$x - y \in S\;\forall y \in \operatorname{spt}\mu_C \text{ and } \forall (x,S) \in \operatorname{spt}C.$$

In particular, choosing T such that $(0,T) \in \operatorname{spt}C$ (such T such that $(0,T)\in\operatorname{spt}C$ (such T exists because $0 \in \operatorname{spt}\mu_C = \pi(\operatorname{spt}C)$), (8) implies $y \in T\ \forall y \in \operatorname{spt}\mu_C$. Thus $\operatorname{spt}\mu_C \subset T$, and hence $C = \theta_0\underline{v}(T)$ by Constancy 4.1.

Thus we have shown that, for $x \in U$ such that 5.6(i), (iii), (iv) hold, each element of $\operatorname{Var}\operatorname{Tan}(V,x)$ has the form $\theta_0\underline{v}(T)$, where T is an n-dimensional subspace of $\mathbb{R}^{n+\ell}$. On the other hand, since we are assuming (5.6(ii)) that $\eta_V^{(x)}$ exists, it follows that for continuous β on $G(n+\ell,n)$

(9) $$\lim_{\rho\downarrow 0}\frac{\int_{G_n(B_\rho(x))}\beta(S)\,dV(y,S)}{\mu_V(B_\rho(x))} = \int_{G(n+\ell,n)}\beta(S)\,d\eta_V^{(x)}(S).$$

Now let $\theta_0\underline{v}(T)$ be any such element of $\operatorname{Var}\operatorname{Tan}(V,x)$ and select $\lambda_j \downarrow 0$ so that $\lim \eta_{x,\lambda_j\#}V = \theta_0\underline{v}(T)$. Then in particular

$$\lim_{j\to\infty} \frac{\int_{G_n(B_1(0))} \beta(S)\, dV_j(y,S)}{\mu_{V_j}(B_1(0))} = \beta(T),$$

and hence (9) gives

$$\beta(T) = \int_{G(n+\ell,n)} \beta(S)\, d\eta_V^{(x)}(S),$$

thus showing that $\theta_0\underline{v}(T)$ is the *unique* element of $\operatorname{Var}\operatorname{Tan}(V,x)$. Thus

$$\lim_{\lambda\downarrow 0} \eta_{x,\lambda\#}V = \theta_0\underline{v}(T),$$

so that T is the tangent space for V at x in the sense of 1.1. This completes the proof. □

The following *compactness theorem* for rectifiable varifolds is now a direct consequence of the Rectifiability 5.5, the Semi-continuity 3.7, and the Compactness Theorem 4.16 of Ch. 1 for Radon measures, and its proof is left to the reader.

5.8 Theorem (Compactness theorem for n-varifolds.) *Suppose $\{V_j\}$ is a sequence of rectifiable n-varifolds in U which are locally bounded first variation in U,*

$$\sup_{j\ge 1}\big(\mu_{V_j}(W) + \|\delta V_j\|(W)\big) < \infty \ \forall W \subset\subset U,$$

and $\Theta^n(\mu_{V_j},x) \ge 1$ on $U \setminus A_j$, where $\mu_{V_j}(A_j \cap W) \to 0$ as $j \to \infty$ $\forall W \subset\subset U$. Then there is a subsequence $\{V_{j'}\}$ and a rectifiable varifold V of locally bounded first variation in U, such that $V_{j'} \to V$ (in the sense of Radon measures on $G_n(U)$), $\Theta^n(\mu_V,x) \ge 1$ for μ_V-a.e. $x \in U$, and $\|\delta V\|(W) \le \liminf_{j\to\infty} \|\delta V_j\|(W)$ for each $W \subset\subset U$.

5.9 Remark: An important additional result (also due to Allard [All72]) is the Integral Compactness Theorem, which asserts that if all the V_j in the above theorem are integer multiplicity, then V is also integer multiplicity. (Notice that in this case the hypothesis $\Theta^n(\mu_{V_j},x) \ge 1$ on $U \setminus A_j$ is automatically satisfied with an A_j such that $\mu_{V_j}(A_j) = 0$.)

Proof that V is integer multiplicity if the V_i are: Let $W \subset\subset U$. We first assert that for μ_V-a.e. $x \in W$ there exists c (depending on x) such that

(1) $$\liminf \|\delta V_i\|(B_\rho(x)) \le c\mu_V(B_\rho(x)),\ \rho < \min\{1, \operatorname{dist}(x,\partial U)\}.$$

Indeed otherwise $\exists$ a set $A \subset W$ with $\mu_V(A) > 0$ such that for each $j \ge 1$ and each $x \in A$ there are $\rho_x > 0$, $i_x \ge 1$ such that $B_{\rho_x}(x) \subset W$ and

$$\mu_V\big(B_{\rho_x}(x)\big) \le j^{-1}\|\delta V_i\|\big(B_{\rho_x}(x)\big),\ i \ge i_x.$$

By the Besicovitch Covering Lemma (§3.12 of Ch. 1) we then have

$$\mu_V(A_i) \le cj^{-1}\|\delta V_\ell\|(W),\ \ell \ge i,$$

where $A_i = \{x \in A : i_x \le i\}$. Thus

$$\mu_V(A_i) \le cj^{-1} \limsup_{\ell\to\infty} \|\delta V_\ell\|(W),$$

and hence $A_i \uparrow A$ as $i \uparrow \infty$ we have

$$\mu_V(A) \le cj^{-1}$$

for some c $(< \infty)$ independent of j. That is, $\mu_V(A) = 0$, a contradiction, and hence (1) holds. Since $\Theta^n(\mu_V,x)$ exists μ_V-a.e. $x \in U$, we in fact have from (1) that for μ_V-a.e. $x \in U$ there is a $c = c(x)$ such that

(2) $$\liminf \|\delta V_i\|(B_\rho(x)) \le c\rho^n,\ 0 < \rho < \min\{1, \operatorname{dist}(x,\partial U)\}.$$

Now since $V = \underline{v}(M,\theta)$, it is also true that for μ_V-a.e. $\xi \in \operatorname{spt}\mu_V$ we have $\eta_{\xi,\lambda\#}V \to \theta_0\underline{v}(P)$ as $\lambda \downarrow 0$, where $P = T_\xi M$ and $\theta_0 = \theta(\xi)$. Then (because $V_i \to V$, and hence $\eta_{\xi,\lambda\#}V_i \to \eta_{\xi,\lambda\#}V$ for each fixed $\lambda > 0$), it follows that for μ_V-a.e. $\xi \in U$ we can select a sequence $\lambda_i \downarrow 0$ such that, with $W_i = \eta_{\xi,\lambda_i\#}V_i$,

(3) $$W_i \to \theta_0\,\underline{v}(P)$$

and (by (2)) for each $R > 0$

(4) $$\|\delta W_i\|(B_R(0)) \to 0.$$

We claim that θ_0 must be an integer for any such ξ; in fact for an arbitrary sequence $\{W_i\}$ of integer multiplicity varifolds in $\mathbb{R}^{n+\ell}$ satisfying (3), (4), we claim that θ_0 always has to be an integer.

To see this, take (without loss of generality) $P = \mathbb{R}^n \times \{0\}$, let q by orthogonal projection onto $(\mathbb{R}^n \times \{0\})^\perp$, and note first that (3) implies

(5) $$p_{\mathbb{R}^n\#}(W_i \llcorner G_n\{x \in \mathbb{R}^{n+\ell} : |q(x)| < \varepsilon\}) \to \theta_0\underline{v}(\mathbb{R}^n)$$

for each fixed $\varepsilon > 0$. However by the mapping formula for varifolds (§1 of Ch. 4), we know that (5) says

(6) $$\underline{v}(\mathbb{R}^n,\psi_i) \to \theta_0\,\underline{v}(\mathbb{R}^n),$$

where

$$\psi_i(x) = \sum_{y\in p^{-1}_{\mathbb{R}^n\times\{0\}}(x)\cap\{z\in\mathbb{R}^{n+\ell}:|q(z)|<\varepsilon\}}\theta_i(y)$$

(θ_i = multiplicity function of W_i, so that ψ_i has values in $\mathbb{Z}\cup\{\infty\}$). Notice that (6) implies in particular that

(7) $$\int_{\mathbb{R}^n} f\psi_i\, d\mathcal{L}^n \to \theta_0 \int_{\mathbb{R}^n} f\, d\mathcal{L}^n \;\forall f\in C^0_c(\mathbb{R}^n).$$

(i.e. measure-theoretic convergence of ψ_i to θ_0.)

Now we claim that there are sets $A_i\subset B_1(0)$ such that

(8) $$\psi_i(x)\le\theta_0+\varepsilon_i\;\forall x\in B_1(0)\setminus A_i,\; \mathcal{L}^n(A_i)\to 0,\;\varepsilon_i\downarrow 0;$$

this will of course (when used in combination with (7)) imply that for any integer $N>\theta_0$, $\max\{\psi_i,N\}$ converges in $L^1(B_1(0))$ to θ_0, and, since $\max\{\psi_i,N\}$ is integer-valued, it then follows that θ_0 is an integer.

On the other hand (8) evidently follows by setting $W=W_i$ in the following lemma, so the proof is complete. □

In this lemma, p, q denote orthogonal projection of $\mathbb{R}^{n+\ell}$ onto $\mathbb{R}^n\times\{0\}\subset\mathbb{R}^{n+\ell}$ and $\{0\}\times\mathbb{R}^k\subset\mathbb{R}^{n+\ell}$ respectively.

5.10 Lemma. *For each $\delta\in(0,1)$, $\Lambda\ge 1$, there is $\varepsilon=\varepsilon(\delta,\Lambda,n)\in(0,\delta^2)$ such that if W is an integer multiplicity varifold in $B_3(0)$ with*

(‡) $$\mu_W(B_3(0))\le\Lambda,\; \|\delta W\|(B_3(0))<\varepsilon^2,\; \int_{B_3(0)}\|p_S-p\|\,dW(y,S)<\varepsilon^2$$

there there is a set $A\subset B^n_1(0)$ such that $\mathcal{L}^n(A)<\delta$ and, $\forall x\in B_1(0)\setminus A$,

$$\sum_{y\in p^{-1}(x)\cap\operatorname{spt}\mu_W\cap\{z:|q(z)|<\varepsilon\}}\Theta^n(\mu_W,y)\le(1+\delta)\frac{\mu_W(B_2(x))}{\omega_n 2^n}+\delta.$$

5.11 Remark: It suffices to prove that for each fixed N there is a $\delta_0=\delta_0(N)\in(0,1)$ such that if $\delta\in(0,\delta_0)$ then $\exists\;\varepsilon=\varepsilon(n,\Lambda,N,\delta)\in(0,\delta^2)$ such that 5.10(‡)implies the existence of $A\subset B^n_1(0)$ with $\mathcal{L}^n(A)<\delta$ and, for $x\in B^n_1(0)\setminus A$ and distinct $y_1,\ldots,y_N\in p^{-1}(x)\cap\operatorname{spt}\mu_W\cap\{z:|q(z)|<\varepsilon\}$,

(‡) $$\sum_{j=1}^N\Theta^n(\mu_W,y_j)\le(1+\delta)\frac{\mu_W(B_2(x))}{\omega_n 2^n}+\delta.$$

Because this firstly implies an *a-priori* bound, depending only on n, k, Λ, on the number N of possible points y_j, and hence the lemma, as originally stated, then follows. (Notice that of course the validity of the lemma for small δ implies its validity for any larger δ.)

Proof of 5.10: By virtue of the above Remark, we need only to prove 5.11(‡). Let $\mu=\mu_W$, and consider the possibility that $y\in B_1(0)$ satisfies the inequalities

(1) $$\delta\|W\|(B_\rho(y))\le\varepsilon\mu(B_\rho(y)),\quad \rho\in(0,1),$$

(2) $$\int_{B_\rho(y)}\|p_S-p\|\,dW(z,S)\le\varepsilon\rho^n,\qquad \rho\in(0,1).$$

Let

$$A_1=\{y\in B_2(0)\cap\operatorname{spt}W:(1)\text{ fails for some }\rho\in(0,1)\}$$
$$A_2=\{y\in B_2(0)\cap\operatorname{spt}W:(2)\text{ fails for some }\rho\in(0,1)\}.$$

Evidently $y\in\operatorname{spt}\mu_W\cap B_2(0)\setminus A_1\Rightarrow$ (by the monotonicity formula 3.3)

(3) $$\frac{\mu(B_\rho(y))}{\omega_n\rho^n}\le e^{\varepsilon}\frac{\mu(B_1(y))}{\omega_n}\le c,\quad 0<\rho<1,$$

($c=c(\Lambda,n)$), while if $y\in A_2\setminus A_1$ we have (using (3))

(4) $$\int_{B_\rho(y)}\|p_S-p\|\,dW(z,S)\ge\varepsilon\rho_y^n\ge c\varepsilon\mu(B_{\rho_y}(y))$$

for some $\rho_y\in(0,1)$. If $y\in A_1$ then

(5) $$\mu(B_{\rho_y}(y))\le\varepsilon^{-1}\|\delta W\|(B_{\rho_y}(y))$$

for some $\rho_y\in(0,1)$.

Since then $\{B_{\rho_y}(y)\}_{y\in A_1\cup A_2}$ covers $A_1\cup A_2$ we deduce from (4), (5) and the Besicovitch Covering (§3.12 of Ch. 1) that

$$\mu(A_1\cup A_2)\le c\varepsilon^{-1}\Big(\int_{B_3(0)}\|p_S-p\|\,dW(a,S)+\|\delta W\|(B_3(0))\Big)$$
$$\le c\varepsilon$$

by the hypotheses on W.

Our aim now is to show 5.11(‡) whenever $x\in B^n_1(0)\setminus p(A_1\cup A_2)$. In view of (6) this will establish the required result (with $A=p(A_1\cup A_2)$). So let $x\in B^n_1(0)\setminus p(A_1\cup A_2)$. In view of the monotonicity formula 3.4 it evidently suffices (by translating and changing scale by a factor of 3/2) to assume that $x=0\in B^n_1(0)\setminus p(A_1\cup A_2)$. We shall subsequently assume this.

We first want to establish the two inequalities, that, for $y\in B^n_1(0)\setminus p(A_1\cup A_2)$ and

$\tau > 0$,

(6) $$\Theta^n(\mu, y) \le e^{\varepsilon\sigma}\frac{\mu(U_\sigma^{2\tau}(y))}{\omega_n\sigma^n} + c\varepsilon\sigma/\tau,\ 0 < \sigma < 1,$$

(7) $$\frac{\mu(U_\sigma^{\tau}(y))}{\omega_n\sigma^n} \le e^{\varepsilon\sigma}\frac{\mu(U_\sigma^{2\tau}(y))}{\omega_n\sigma^n} + c\varepsilon\sigma/\tau,\ 0 < \sigma < \rho \le 1,$$

where
$$U_\sigma^\tau(y) = B_\sigma(y) \cap \{z \in \mathbb{R}^{n+\ell} : |q(z-y)| < \tau\}.$$

Indeed these two inequalities follow directly from 3.3 and 3.5. For example to establish (6) we note first that 3.3 gives (6) directly if $\tau \ge \sigma$, while if $\tau < \sigma$ then we first use 3.3 to give $\Theta^n(\mu, y) \le e^{\varepsilon\tau}\frac{\mu(B_\tau(y))}{\omega_n\tau^n}$ and then use 3.5 with h of the form $h(z) = f(|q(z-y)|)$, $f(t) \equiv 1$ for $t < \tau$ and $f(t) \equiv 0$ for $t > 2\tau$.

Since $\left|\nabla^S f(|q(z-y)|)\right| \le f'(|q(z-y)|)|p_S - p|$ (Cf. the computation in ?? of Ch. 4 we then deduce (by integrating in 3.5 from τ to σ and using (3))

$$\frac{\mu(B_\tau(y))}{\omega_n\tau^n} \le \frac{\mu(U_\sigma^{2\tau}(y))}{\omega_n\sigma^n} + c\varepsilon\sigma/\tau.$$

(7) is proved by simply integrating 3.5 from σ to ρ (and using (3)).

Our aim now is to use (6) and (7) to establish

(8) $$\sum_{j=1}^N \frac{\mu(U_\sigma^\tau(y_j))}{\omega_n\sigma^n} \le (1 + c\delta^2)\frac{\mu(B_2(0))}{\omega_n 2^n} + c\delta^2$$

with $c = c(n, k, N, \Lambda)$, provided $2\delta^2\sigma \le \tau \le \frac{1}{4}\min_{j\ne\ell}|y_j - y_\ell|$, $y_j \in \operatorname{spt}\mu \cap p^{-1}(0) \cap \{z : |q(z)| < \varepsilon\}$, $0 \notin p(A_1 \cup A_2)$. (In view of (6) this will prove the required result 5.11 (‡) for suitable $\delta_0(N)$.)

We proceed by induction on N. $N = 1$ trivially follows from (7) by noting that $U_\rho^{2\tau}(y_1) \subset B_\rho(y_1)$ (by definition of $U_\rho^{2\tau}(y_1)$) and then using the monotonicity 3.3 together with the fact that $|y_1| < \varepsilon$. Thus assume $N \ge 2$ and that (8) has been established with any $M < N$ in place of N.

Let $y_1, \ldots, y_N$ be as in (8), and choose $\rho \in [\sigma, 1)$ such that $\min_{j\ne\ell}|q(y_j) - q(y_\ell)|$ $(= \min_{j\ne\ell}|y_j - y_\ell|) = 4\delta^2\rho$, and set $\widetilde\tau = 2\delta^2\rho\ (\ge 2\tau)$. Then

$$\begin{aligned}\frac{\mu(U_\sigma^\tau(y_j))}{\sigma^n} &\le \frac{\mu(U_\sigma^{\frac12\widetilde\tau}(y_j))}{\sigma^n}\\ &\le e^{\varepsilon\rho}\frac{\mu(U_\rho^{\widetilde\tau}(y_j))}{\rho^n} + c\varepsilon \quad \text{(by (7))},\end{aligned}$$

$c = c(n, k, \delta)$. Now since $\widetilde\tau = \frac12\min_{j\ne\ell}|q(y_j) - q(y_\ell)|$ we can select $\{z_1, \ldots, z_Q\} \subset \{y_1, \ldots, y_N\}$ $(Q \le N-1)$ and $\widetilde\tau \le c\widetilde\tau$ such that $\widehat\tau \ge 3\delta^2\rho$ and

$$\cup_{j=1}^N U_\rho^{\widetilde\tau}(y_j) \subset \cup_{\ell=1}^Q U_{\rho(1+c\delta^2)}^{\widehat\tau}(z_\ell),$$

where $c = c(N)$, and such that $\widehat\tau \le \frac14\min_{i\ne j}|z_i - z_j|$. Since $c\delta^2 < 1/2$ for $\delta < \delta_0(N)$ (if $\delta_0(N)$ is chosen suitably) we then $\widehat\tau \ge 2\delta^2\widetilde\rho$ and

$$\sum_{j=1}^N \frac{\mu(U_\rho^{\widetilde\tau}(y_j))}{\rho^n} \le (1 + c\delta^2)\sum_{j=1}^Q \frac{\mu(U_{\widetilde\rho}^{\widehat\tau}(z_j))}{\widetilde\rho^n},$$

where $\widetilde\rho = (1 + c\delta^2)\rho$ and $c = c(N)$. Since $Q \le N - 1$, the required result then follows by induction (choosing ε appropriately). □

Appendix A

A General Regularity Theorem

We here prove a useful general regularity theorem, which is essentially an abstraction of the "dimension reducing" argument of [Fed70]. There are a number of important applications of this general theorem in the text.

Let $P \geq n \geq 2$ and let $\mathcal{F}$ be a collection of functions $\varphi = (\varphi^1, \dots, \varphi^Q) : \mathbb{R}^P \to \mathbb{R}^Q$ ($Q = 1$ is an important case) such that each φ^j is locally $\mathcal{H}^n$-integrable on $\mathbb{R}^P$. For $\varphi \in \mathcal{F}$, $y \in \mathbb{R}^P$ and $\lambda > 0$ we let $\varphi_{y,\lambda}$ be defined by

$$\varphi_{y,\lambda}(x) = \varphi(y + \lambda x), \; x \in \mathbb{R}^P.$$

Also, for $\varphi \in \mathcal{F}$ and a given sequence $\{\varphi_k\} \subset \mathcal{F}$ we write $\varphi_k \rightharpoonup \varphi$ if $\int \varphi_k f \, d\mathcal{H}^n \to \int \varphi f \, d\mathcal{H}^n$ (in $\mathbb{R}^Q$) for each given $f \in C_c^0(\mathbb{R}^P)$.

We subsequently make the following 3 special assumptions concerning $\mathcal{F}$:

A.1 (Closure under appropriate scaling and translation): If $|y| \leq 1 - \lambda$, $0 < \lambda < 1$, and if $\varphi \in \mathcal{F}$, then $\varphi_{y,\lambda} \in \mathcal{F}$.

A.2 (Existence of homogeneous degree zero "tangent functions"): If $|y| < 1$, if $\{\lambda_k\} \downarrow 0$ and if $\varphi \in \mathcal{F}$, then there is a subsequence $\{\lambda_{k'}\}$ and $\psi \in \mathcal{F}$ such that $\varphi_{y,\lambda_{k'}} \rightharpoonup \psi$ and $\psi_{0,\lambda} = \psi$ for each $\lambda > 0$.

A.3 ("Singular set" hypotheses): We assume there is a map

$$\operatorname{sing} : \mathcal{F} \to \mathcal{C} \; (= \text{ set of closed subsets of } \mathbb{R}^P)$$

such that:

(1) $\operatorname{sing}\varphi = \emptyset$ if $\varphi \in \mathcal{F}$ is a constant multiple of the indicator function of an n-dimensional subspace of $\mathbb{R}^P$,

(2) If $|y| \leq 1 - \lambda$, $0 < \lambda < 1$, then $\operatorname{sing}\varphi_{y,\lambda} = \lambda^{-1}(\operatorname{sing}\varphi - y)$,

(3) If $\varphi, \varphi_k \in \mathcal{F}$ with $\varphi_k \rightharpoonup \varphi$, then for each $\varepsilon > 0$ there is a $k(\varepsilon)$ such that

$$B_1(0) \cap \operatorname{sing}\varphi_k \subset \{x \in \mathbb{R}^P : \operatorname{dist}(\operatorname{sing}\varphi, x) < \varepsilon\} \ \forall k \geq k(\varepsilon).$$

We can now state the main result of this section:

A.4 Theorem. *Subject to the notation and assumptions* A.1, A.2, A.3 *above we have*

$$(\ddagger) \qquad \dim(B_1(0) \cap \operatorname{sing}\varphi) \leq n-1 \ \forall \varphi \in \mathcal{F}.$$

(Here "dim" is Hausdorff dimension, i.e. (‡) means $\mathcal{H}^{n-1+\alpha}(\operatorname{sing}\varphi) = 0 \ \forall \alpha > 0$.)

In fact either $\operatorname{sing}\varphi \cap B_1(0) = \emptyset$ *for every* $\varphi \in \mathcal{F}$ *or else there is an integer* $d \in [0, n-1]$ *such that*

$$\dim \operatorname{sing}\varphi \cap B_1(0) \leq d \ \forall \varphi \in \mathcal{F}$$

and such that there is some $\psi \in \mathcal{F}$ *and a* d*-dimensional subspace* $L \subset \mathbb{R}^P$ *with*

$$(\ddagger\ddagger) \qquad \operatorname{sing}\psi = L \text{ and } \psi_{y,\lambda} = \psi \ \ \forall y \in L, \ \lambda > 0.$$

If $d = 0$ *then* $\operatorname{sing}\varphi \cap B_\rho(0)$ *is finite for each* $\varphi \in \mathcal{F}$ *and each* $\rho < 1$.

A.5 Remark: One readily checks that if L is an n-dimensional subspace of $\mathbb{R}^P$ and $\psi \in \mathcal{F}$ satisfies A.4 (‡‡), then ψ is exactly a constant multiple of the indicator function of L (hence $\operatorname{sing}\psi = \emptyset$ by A.3(1)); otherwise we would have $P > n$ and $\psi \equiv$ const. $\neq 0$ on some $(n+1)$-dimensional half-space, thus contradicting the fact that ψ is locally $\mathcal{H}^n$-integrable on $\mathbb{R}^P$.

Proof of A.4: Assume $\operatorname{sing}\varphi \cap B_1(0) \neq \emptyset$ for some $\varphi \in \mathcal{F}$, and let $d = \sup\{\dim L :$ L is a d-dimensional subspace of $\mathbb{R}^P$ and there is $\varphi \in \mathcal{F}$ with $\operatorname{sing}\varphi \neq \emptyset$ and $\varphi_{y,\lambda} = \varphi \ \forall y \in L, \lambda > 0\}$. Then by A.5 we have $d \leq n-1$.

For a given $\varphi \in \mathcal{F}$ and $y \in B_1(0)$ we let $T(\varphi, y)$ be the set of $\psi \in \mathcal{F}$ with $\psi_{0,\lambda} = \psi \ \forall \lambda > 0$ and with $\lim \varphi_{y,\lambda_k} = \psi$ for some sequence $\lambda_k \downarrow 0$. ($T(\varphi, y) \neq \emptyset$ by assumption A.2.)

Let $\ell \geq 0$ and let

$$\mathcal{F}^\ell = \left\{\varphi \in \mathcal{F} : \mathcal{H}^\ell(\operatorname{sing}\varphi \cap B_1(0)) > 0\right\}.$$

Our first task is to prove the implication

$$(1) \qquad \varphi \in \mathcal{F}^\ell \Rightarrow \exists \psi \in T(\varphi, x) \cap \mathcal{F}^\ell$$

for $\mathcal{H}^\ell$-a.e. $x \in \operatorname{sing}\varphi \cap B_1(0)$.

To see this, let $\mathcal{H}^\ell_\delta$ be the "size δ approximation" of $\mathcal{H}^\ell$ as described in §2 of Ch. 1 and recall $\mathcal{H}^\ell(A) > 0 \iff \mathcal{H}^\ell_\infty(A) > 0$, so that

$$\mathcal{F}^\ell = \left\{\varphi \in \mathcal{F} : \mathcal{H}^\ell_\infty(\operatorname{sing}\varphi \cap B_1(0)) > 0\right\}.$$

Also note that (by 3.6(2) of Ch. 1), for any bounded subset A of $\mathbb{R}^P$,

$$\mathcal{H}^\ell_\infty(A) > 0 \Rightarrow \Theta^{*n}(\mathcal{H}^\ell_\infty \llcorner A, x) > 0 \quad \text{for} \quad \mathcal{H}^\ell\text{-a.e. } x \in A.$$

Thus we see that if $\varphi \in \mathcal{F}^\ell$ then for $\mathcal{H}^\ell$-a.e. $x \in \operatorname{sing}\varphi \cap B_1(0)$ we have

$$\Theta^{*\ell}(\mathcal{H}^\ell_\infty \llcorner \operatorname{sing}\varphi, x) > 0.$$

For such x we thus have a sequence $\lambda_k \downarrow 0$ such that

$$(2) \qquad \lim_{k\to\infty} \frac{\mathcal{H}^\ell_\infty(\operatorname{sing}\varphi \cap B_{\lambda_k}(x))}{\lambda_k^\ell} > 0,$$

and by assumption A.2 there is a subsequence $\{\lambda_{k'}\}$ such that $\varphi_{x,\lambda_{k'}} \rightharpoonup \psi \in T(\varphi, x)$. If now $\mathcal{H}^\ell_\infty(\operatorname{sing}\psi) = 0$, then for any $\varepsilon > 0$ we could find open balls $\{B_{\rho_j}(x_j)\}$ such that

$$\operatorname{sing}\psi \subset \cup_j B_{\rho_j}(x_j)$$

and

$$(3) \qquad \textstyle\sum_j \omega_\ell \rho_j^\ell < \varepsilon$$

(be definition of $\mathcal{H}^\ell_\infty$). Now (2) in particular implies that $K \equiv B_1(0) \setminus \cup_j B_{\rho_j}(x_j)$ is a compact set with positive distance from $\operatorname{sing}\psi$. Hence by assumption A.3(3) we have

$$\operatorname{sing}\varphi_{x,\lambda_{k'}} \cap B_1(0) \subset \cup_j B_{\rho_j}(x_j)$$

for all sufficiently large k, and hence by (3)

$$\mathcal{H}^\ell_\infty(\operatorname{sing}\varphi_{x,\lambda_{k'}} \cap B_1(0)) < \varepsilon, \ k \geq k(\varepsilon).$$

Thus since $\lambda_k^{-1}(\operatorname{sing}\varphi - x) = \operatorname{sing}\varphi_{x,\lambda_k}$ (by A.3(2)) we have

$$\lambda_{k'}^{-\ell}\mathcal{H}^\ell_\infty(\operatorname{sing}\varphi \cap B_{\lambda_{k'}}(x)) < \varepsilon$$

for all sufficiently large k, thus a contradiction for $\varepsilon < \lim_{k\to\infty} \lambda_k^{-\ell}\mathcal{H}^\ell_\infty(\operatorname{sing}\varphi \cap B_{\lambda_k}(x))$. (Such ε can be chosen by (2).)

We have therefore established the general implication (1). From now on take $\ell > d-1$ so that $\mathcal{F}^\ell \neq \emptyset$ (which is automatic for $\ell \leq d$ by definition of d). By (1) there is $\varphi \in \mathcal{F}^\ell$ with $\varphi_{0,\lambda} = \varphi \ \forall \lambda > 0$. Suppose also that there is a k-dimensional subspace ($k \geq 0$) S of $\mathbb{R}^P$ such that $\varphi_{y,\lambda} = \varphi \ \forall y \in S, \lambda > 0$. (Notice of course this is no additional restriction for φ in case $k = 0$.) Now if $k \geq d+1$ then, by definition of d, we can assert $\operatorname{sing}\varphi = \emptyset$, thus contradicting the fact that $\varphi \in \mathcal{F}^\ell$. Therefore $0 \le k \le d$, and if $k \le d-1$ ($< \ell$), then $\mathcal{H}^\ell(S) = 0$ and in particular

(4) $$\exists x \in B_1(0) \cap \operatorname{sing}\varphi \setminus S.$$

But by A.2 we can choose $\psi \in T(\varphi, x)$. Since $\psi = \lim \varphi_{x,\lambda_j}$ for some sequence $\lambda_j \downarrow 0$, we evidently have (since $\varphi_{y+x,\lambda} = \varphi_{x,\lambda}$ $\forall y \in S$, $\lambda > 0$)

$$\psi_{y,1} = \lim \varphi_{y+x,\lambda_j} = \lim \varphi_{x,\lambda_j} = \psi \ \forall y \in S$$

and

$$\psi_{\beta x,1} = \lim \varphi_{x+\lambda_j \beta x,\lambda_j} = \psi \ \forall \beta \in \mathbb{R}.$$

(All limits in the weak sense described at the beginning of the section.) Thus $\psi_{z,\lambda} = \psi$ for each $\lambda > 0$ and each z in the $(k+1)$-dimensional subspace T of $\mathbb{R}^P$ spanned by S and x. $\operatorname{sing}\psi \ne \emptyset$ (by A.3(3)), hence by induction on k we can take $k = d-1$; i.e. $\dim T = d$, and hence $\operatorname{sing}\psi \subset T$ by A.3(2). On the other hand if $\exists \tilde{x} \in \operatorname{sing}\psi \setminus T$ then we can repeat the above argument (beginning at (4)) with T in place of S and ψ in place of φ. This would then give a $(d+1)$-dimensional subspace $\widetilde{T}$ and a $\widetilde{\psi} \in \mathcal{F}$ with $\operatorname{sing}\widetilde{\psi} \supset \widetilde{T}$, thus contradicting the definition of d. Therefore $\operatorname{sing}\varphi = T$. Furthermore if $\ell > d$ then the above induction works up to $k = d$ and again therefore we would have a contradiction. Thus $\dim(B_1(0) \cap \operatorname{sing}\varphi) \le d$ $\forall \varphi \in \mathcal{F}$.

Finally to prove the last claim of the theorem, we suppose that $d = 0$. Then we have already established that

(5) $$\mathcal{H}^\alpha(\operatorname{sing}\varphi \cap B_1(0)) = 0 \ \forall \alpha > 0, \ \varphi \in \mathcal{F}.$$

If $\operatorname{sing}\varphi \cap B_\rho(0)$ is not finite, then we select $x \in B_\rho(0)$ such that $x = \lim x_k$ for some sequence $x_k \in \operatorname{sing}\varphi \cap B_1(0) \setminus \{x\}$. Then letting $\lambda_k = 2|x_k - x|$ we see from A.3(2) that there is a subsequence $\{\lambda_{k'}\}$ with $\varphi_{x,\lambda_{k'}} \rightharpoonup \psi \in T(\varphi, x)$ and $(x_{k'} - x)/|x_{k'} - x| \to \xi \in \partial B_1(0)$. Now by A.3(2), (3) we know that $\{\xi/2\} \cap \{0\} \subset \operatorname{sing}\psi$ and, since $\psi_{0,\lambda} = \psi$, this (together with A.3(2)) gives $L_\xi \subset \operatorname{sing}\psi$ where L_ξ is the ray determined by 0 and ξ. Then $\mathcal{H}^1(\operatorname{sing}\psi \cap B_1(0)) > 0$, thus contradicting (5), because $\psi \in \mathcal{F}$. □

Appendix B

Non-existence of Stable Minimal Hypercones, *n* ≤ 6

Here we describe J. Simons [Sim68] result on non-existence in $\mathbb{R}^{n+1}$ of n-dimensional stable minimal cones (previously established in case $n = 2, 3$ by Fleming [Fle62] and Almgren [Alm66] respectively). The proof here follows essential Schoen-Simon-Yau [SSY75], which is a slight variant of the original proof in [Sim68].

Suppose to begin that $C \in \mathcal{D}_n(\mathbb{R}^{n+1})$ is a cone ($\eta_{0,\lambda\#}C = C$) and C is integer multiplicity with $\partial C = 0$. If $\operatorname{sing} C \subset \{0\}$ and if C is minimizing in $\mathbb{R}^{n+1}$ then, writing $M = \operatorname{spt} C \setminus \{0\}$ and taking M_t as in §5 of Ch. 2, we have $\frac{d}{dt}\mathcal{H}^n(M_t)\big|_{t=0} = 0$ and $\frac{d^2}{dt^2}\mathcal{H}^n(M_t)\big|_{t=0} \ge 0$. (This is clear because in fact $\mathcal{H}^n(M_t)$ takes its minimum value at $t = 0$, by virtue of our assumption that C is minimizing.) Notice that M is orientable, with orientation induced from C, and hence in particular we can deduce from 5.12 of Ch. 2 that

B.1 $$\int_M \left(|\nabla^M \zeta|^2 - \zeta^2 |A|^2\right) d\mathcal{H}^n \ge 0$$

for every $\zeta \in C_c^1(M)$ (notice $0 \notin M$, so such ζ vanish in a neighborhood of 0). Here A is the second fundamental form of M and $|A|$ is its length, as described in § 4 of Ch. 2 and 5.12 of Ch. 2.

The main result we need is given in the following theorem.

B.2 Theorem. *Suppose $2 \le n \le 6$ and M is an n-dimensional cone embedded in $\mathbb{R}^{n+1}$ with zero mean curvature (see §4 of Ch. 2) and with $\overline{M} \setminus M = \{0\}$, and suppose that M is stable in the sense that B.1 holds. Then $\overline{M}$ is a hyperplane.*

As explained above, the hypotheses are in particular satisfied if $M = \operatorname{spt} C \setminus \{0\}$, with $C \in \mathcal{D}_n(\mathbb{R}^{n+1})$ a minimizing cone with $\partial C = 0$ and $\operatorname{sing} C \subset \{0\}$.)

B.3 Remark: B.2 is false for $n = 7$; J. Simons [Sim68] was the first to point out that the cone $M = \left\{(x^1, \ldots, x^8) \in \mathbb{R}^8 : \sum_{i=1}^4 (x^i)^2 = \sum_{i=5}^8 (x^i)^2\right\}$ is a stable minimal cone. (Notice that M is the cone over the compact manifold $(\frac{1}{\sqrt{2}}\mathbb{S}^3) \times (\frac{1}{\sqrt{2}}\mathbb{S}^3) \subset \mathbb{S}^7 \subset \mathbb{R}^8$.) The fact that the mean curvature of M is zero is checked by direct computation. The fact that M is actually *stable* is checked as follows. First, by direct computation one checks that the second fundamental form A of M satisfies $|A|^2 = 6/|x|^2$.

On the other hand for a stationary hypersurface $M \subset \mathbb{R}^{n+1}$ the first variation formula 5.3 of Ch. 2 says $\int \operatorname{div}_M X \, d\mathcal{H}^n = 0$ if $\operatorname{spt}|X|$ is a compact subset of M. Taking $X_x = (\zeta^2/r^2)x$, $\zeta \in C_c^\infty(M)$, $r = |x|$, and computing as in §4 of Ch. 4, we get

$$(n-2)\int_M (\zeta^2/r^2)\, d\mathcal{H}^n = -2\int_M \zeta r^{-2} x \cdot \nabla^M \zeta \, d\mathcal{H}^n.$$

Using the Schwarz inequality on the right we get

$$\frac{(n-2)^2}{4}\int_M (\zeta^2/r^2)\, d\mathcal{H}^n \le \int_M |\nabla^M \zeta|^2 \mathcal{H}^n.$$

Thus we have stability for M (in the sense of B.1) whenever A satisfies $|x|^2|A|^2 \le (n-2)^2/4$.

For the example above we have $n = 7$ and $|x|^2|A|^2 = 6$, so that this inequality is satisfied, and the cone over $\mathbb{S}^3 \times \mathbb{S}^3$ is stable as claimed. (Similarly the cone over $\mathbb{S}^q \times \mathbb{S}^q$ is stable for $q \ge 3$; i.e. when the dimension of the cone is ≥ 7.)

Before giving the proof of B.2 we need to derive the identity of J. Simons for the Laplacian of the length of the second fundamental form of a hypersurface (B.8 below).

The simple derivation here assumes the reader's familiarity with basic Riemannian geometry. (A completely elementary derivation, assuming no such background, is described in [Giu84].)

For the moment let M be an arbitrary hypersurface in $\mathbb{R}^{n+1}$ (M not necessarily a cone, and not necessarily having zero mean curvature).

Let $\tau_1, \ldots, \tau_n$ be a locally defined family of smooth vector fields which, together with the unit normal ν of M, define an orthonormal basis for $\mathbb{R}^{n+1}$ at all points in some region of M.

The second fundamental form of M relative to the unit normal ν is the tensor $A = h_{ij}\tau_i \otimes \tau_j$, where $h_{ij} = \langle D_{\tau_j}\nu, \tau_i\rangle$. (Cf. §4 of Ch. 2.) Recall (see 4.26 of Ch. 2) that

B.4 $$h_{ij} = h_{ji},$$

and, since the Riemann tensor of the ambient space $\mathbb{R}^{n+1}$ is zero, we have the Codazzi equations

B.5 $$h_{ij,k} = h_{ik,j}, \; i, j, k \in \{1, \ldots, n\}.$$

Here $h_{ij,k}$ denotes the covariant derivative of A with respect to τ_k; that is, $h_{ij,k}$ are such that $\nabla_{\tau_k} A = h_{ij,k}\tau_i \otimes \tau_j$.

We also have the Gauss curvature equations

B.6 $$R_{ijk\ell} = h_{i\ell}h_{jk} - h_{ik}h_{j\ell},$$

where $R = R_{ijk\ell}\tau_i \otimes \tau_j \otimes \tau_k \otimes \tau_\ell$ is the Riemann curvature tensor of M, and where we use the sign convention such that R_{ijji} ($i \ne j$) are sectional curvatures of M ($= +1$, if $M = \mathbb{S}^n$).

From the properties of R (in fact essentially by definition of R) we also have, for any 2-tensor $a_{ij}\tau_i \otimes \tau_j$,

$$a_{ij,k\ell} = a_{ij,\ell k} + a_{im}R_{mj\ell k} + a_{mj}R_{mi\ell k}$$

(where $a_{ij,k\ell}$ means $a_{ij,k,\ell}$—i.e. the covariant derivative with respect to τ_ℓ of the tensor $a_{ij,k}\tau_i \otimes \tau_j \otimes \tau_k$). In particular

B.7 $$\begin{aligned} h_{ij,k\ell} &= h_{ij,\ell k} + h_{im}R_{mj\ell k} + h_{mj}R_{mi\ell k} \\ &= h_{ij,\ell k} + h_{im}[h_{m\ell}h_{jk} - h_{mk}h_{j\ell}] - h_{mj}[h_{i\ell}h_{mk} - h_{ik}h_{m\ell}] \end{aligned}$$

by B.6, where, here and subsequently, repeated indices are summed from 1 to n.

B.8 Lemma. *In the notation above,*

$$\Delta_M\left(\frac{1}{2}|A|^2\right) = \textstyle\sum_{i,j,k} h_{ij,k}^2 - |A|^4 + h_{ij}H_{,ij} + H h_{mi}h_{mj}h_{ij},$$

where $H = h_{kk} = $ trace A.

Proof: We first compute $h_{ij,kk}$:

$$\begin{aligned} h_{ij,kk} &= h_{ik,jk} \quad \text{(by B.5)} \\ &= h_{ki,jk} \quad \text{(by B.4)} \\ &= h_{ki,kj} + h_k m[h_{mj}h_{ik} - h_{mk}h_{ij}] \\ &\quad - h_{mi}[h_{kj}h_{mk} - h_{kk}h_{mj}] \quad \text{(by B.7)} \\ &= h_{ki,kj} - (\textstyle\sum_{m,k} h_{mk}^2)h_{ij} + h_{kk}h_{mi}h_{mj} \\ &= h_{kk,ij} - (\textstyle\sum_{m,k} h_{mk}^2)h_{ij} + h_{kk}h_{mi}h_{mj} \quad \text{(by B.5)} \end{aligned}$$

Now multiplying by h_{ij} we then get (since $h_{ij}h_{ij,kk} = \frac{1}{2}(\sum_{i,j} h_{ij}^2)_{,kk} - \sum_{i,j,k} h_{ij,k}^2$)

B.9 $$\frac{1}{2}(\textstyle\sum_{i,j} h_{ij}^2)_{,kk} = \sum_{i,j,k} h_{ij,k}^2 = (\sum_{i,j} h_{ij}^2)^2 + h_{ij}H_{,ij} + Hh_{mi}h_{mj}h_{ij},$$

which is the required identity.

We now want to examine carefully the term $\sum_{i,j,k} h_{ij,k}^2$ appearing in the identity of B.8 in case M is a cone with vertex at 0 (i.e. $\eta_{0,\lambda}M = M \ \forall \lambda > 0$). In particular we want to compare $\sum_{i,j,k} h_{ij,k}^2$ with $|\nabla^M |A||^2$ in this case. Since $|\nabla^M|A||^2 = \sum_{k=1}^n |A|^{-2}(h_{ij}h_{ij,k})^2$, we look at the difference

B.10 $$D \equiv \textstyle\sum_{i,j,k} h_{ij,k}^2 - \sum_{k=1}^n |A|^{-2}(h_{ij}h_{ij,k})^2.$$

B.11 Lemma. *If M is a cone (not necessarily minimal) the quantity D defined in* B.10 *satisfies*

$$D(x) \ge 2|x|^{-2}|A(x)|^2, \ x \in M.$$

Proof: Let $x \in M$ and select the frame $\tau_1, \dots, \tau_n$ so that τ_n is radial ($x/|x|$) along the ray ℓ_x through x, and so (as vectors in $\mathbb{R}^{n+1}$) $\tau_1, \dots, \tau_n$ are constant along ℓ_x. Then

(1) $$h_{nj} = h_{jn} = 0 \quad \text{on } \ell_x, \ j = 1, \dots, n,$$

and (since $h_{ij}(\lambda x) = \lambda^{-1}h_{ij}(x)$, $\lambda > 0$)

(2) $$h_{ij,n} = -r^{-1}h_{ij} \quad \text{on } \ell_x.$$

Rearranging the expression for D, we have

$$D = \frac{1}{2}\textstyle\sum_{k=1}^n \sum_{i,j,r,s=1}^n |A|^{-2}(h_{rs}h_{ij,k} - h_{ij}h_{rs,k})^2,$$

as one easily checks by expanding the square on the right. Now since

$$\textstyle\sum_{i,j,r,s=1}^n (h_{rs}h_{ij,k} - h_{ij}h_{rs,k})^2 \ge 4\sum_{\substack{i,j,r=1\\ s=n}}^{n-1} (h_{rs}h_{ij,k} - h_{ij}h_{rs,k})^2,$$

we thus have

$$D \ge 2|A|^{-2}\textstyle\sum_{k=1}^n \sum_{i,j,r=1}^{n-1} (h_{ij}h_{rn,k})^2.$$

By the Codazzi equations B.5 and (2) this gives

$$\begin{aligned} D &\ge 2r^{-2}|A|^{-2}\textstyle\sum_{k=1}^n \sum_{i,j,r=1}^{n-1} h_{ij}^2 h_{rk}^2 \\ &= 2r^{-2}|A|^{-2}|A|^4 \quad \text{(by (1))} \\ &= 2r^{-2}|A|^2, \end{aligned}$$

as required. □

Proof of B.2: Notice that so far we have not used the minimality of M (i.e. we have not used $H(= h_{kk}) = 0$). We now do set $H = 0$ in the above computations, thus giving (by B.8, B.11)

(1) $$\Delta_M\left(\frac{1}{2}|A|^2\right) + |A|^4 \ge 2r^{-2}|A|^2 + |\nabla|A||^2$$

for the minimal cone M. (Notice that $|A|$ is Lipschitz, and hence $|\nabla|A||$ makes sense $\mathcal{H}^n$-a.e. in M.)

Our aim now is to use (1) in combination with the stability inequality B.1 to get a contradiction in case $2 \le n \le 6$.

Specifically, replace ζ by $\zeta|A|$ in B.1. This gives

(2) $$\begin{aligned} \int_M \zeta^2|A|^4 &\le \int_M |\nabla(\zeta|A|)|^2 \\ &= \int_M (|\nabla\zeta|^2|A|^2 + \zeta^2|\nabla|A||^2) + 2\int_M \zeta|A|\nabla\zeta\cdot\nabla|A|. \end{aligned}$$

Now

$$\begin{aligned} 2\int_M \zeta|A|\nabla\zeta\cdot\nabla|A| &= 2\int_M \zeta\nabla\zeta\cdot\nabla\left(\tfrac{1}{2}|A|^2\right) \\ &= \int_M (\nabla\zeta^2)\cdot\nabla\left(\tfrac{1}{2}|A|^2\right) \\ &= -\int_M \zeta^2\Delta_M\left(\tfrac{1}{2}|A|^2\right) \\ &\le \int_M \left(|A|^5\zeta^2 - 2r^{-2}\zeta^2|A|^2 + \zeta^2|\nabla|A||^2\right) \quad \text{by (1),} \end{aligned}$$

and hence (2) gives

$$2\int_M r^{-2}\zeta^2|A|^2 \le \int_M |A|^2|\nabla\zeta|^2 \quad \forall \zeta \in C_c^1(M). \tag{3}$$

Now we claim that (3) is valid even if ζ does not have compact support on M, provided that ζ is locally Lipschitz and

$$\int_M r^{-2}\zeta^2|A|^2 < \infty. \tag{4}$$

(This is proved by applying (3) with $\zeta\gamma_\varepsilon$ in place of ζ, where γ_ε is such that $\gamma_\varepsilon(x) \equiv 1$ for $|x| \in (\varepsilon, \varepsilon^{-1})$, $|\nabla\gamma_\varepsilon(x)| \le 3/|x|$ for all x, $\gamma_\varepsilon(x) = 0$ for $|x| < \varepsilon/2$ or $|x| > 2\varepsilon^{-1}$, and $0 \le \gamma_\varepsilon \le 1$ everywhere, then letting $\varepsilon \downarrow 0$ and using (4).)

Since M is a cone we can write

$$\int_M \varphi(x)\, d\mathcal{H}^n(x) = \int_0^\infty r^{n-1}\int_\Sigma \varphi(r\omega)\, d\mathcal{H}^{n-1}(\omega) dr \tag{5}$$

for any non-negative continuous φ on M, where $\Sigma = M \cap \mathbb{S}^n$ is a compact $(n-1)$-dimensional submanifold. Since $|A(x)|^2 = r^{-2}|A(x/|x|)|^2$, we can now use (5) to check that $\zeta = r^{1+\varepsilon}r_1^{1-n/2-2\varepsilon}$, $r_1 = \max\{1, r\}$, is a valid choice to ensure (4), hence we may use this choice in (3). This is easily seen to give

$$\begin{aligned}2\int_M r^{2\varepsilon}r_1^{2-n-4\varepsilon}|A|^2 \le (\tfrac{n}{2} - 2 + \varepsilon)^2 &\int_{M\cap\{r>1\}} |A|^2 r^{2-n-2\varepsilon} \\ &+ (1+\varepsilon)^2\int_{M\cap\{r<1\}} |A|^2 r^{2\varepsilon} < \infty.\end{aligned} \tag{6}$$

For $2 \le n \le 6$ we can choose ε such that $(\frac{n}{2} - 2 + \varepsilon)^2 < 2$ and $(1+\varepsilon)^2 < 2$, hence (6) gives $|A|^2 \equiv 0$ on M as required. □

Bibliography

[All72] William K. Allard, *On the first variation of a varifold*, Ann. of Math. (2) **95** (1972), 417–491. MR MR0307015 (46 #6136)

[Alm66] F. J. Almgren, Jr., *Some interior regularity theorems for minimal surfaces and an extension of Bernstein's theorem*, Ann. of Math. (2) **84** (1966), 277–292. MR MR0200816 (34 #702)

[Alm68] ______, *Existence and regularity almost everywhere of solutions to elliptic variational problems among surfaces of varying topological type and singularity structure*, Ann. of Math. (2) **87** (1968), 321–391. MR MR0225243 (37 #837)

[Alm84] ______, *$\underline{\underline{Q}}$- valued functions minimizing dirichlet's integral and the regularity of area minimizing rectifiable currents up to codimension* 2, Preprint, Princeton University, 1984.

[Bes28] A. S. Besicovitch, *On the fundamental geometrical properties of linearly measurable plane sets of points*, Math. Ann. **98** (1928), no. 1, 422–464. MR MR1512414

[Bes38] ______, *On the fundamental geometrical properties of linearly measurable plane sets of points (II)*, Math. Ann. **115** (1938), no. 1, 296–329. MR MR1513189

[Bes39] ______, *On the fundamental geometrical properties of linearly measurable plane sets of points (III)*, Math. Ann. **116** (1939), no. 1, 349–357. MR MR1513231

[DG61] Ennio De Giorgi, *Frontiere orientate di misura minima*, Seminario di Matematica della Scuola Normale Superiore di Pisa, 1960-61, Editrice Tecnico Scientifica, Pisa, 1961. MR MR0179651 (31 #3897)

[Dug86] J. P. Duggan, *Regularity theorems for varifolds with mean curvature*, Indiana Univ. Math. J. **35** (1986), no. 1, 117–144. MR MR825631 (87e:49065)

[EG92] C. Evans and R. Gariepy, *Measure theory and fine properties of functions*, Studies in Advanced Mathematics, CRC Press, 1992. MR MR0257325 (41 #1976)

[Fed69] Herbert Federer, *Geometric measure theory*, Die Grundlehren der mathematischen Wissenschaften, Band 153, Springer-Verlag New York Inc., New York, 1969. MR MR0257325 (41 #1976)

[Fed70] ———, *The singular sets of area minimizing rectifiable currents with codimension one and of area minimizing flat chains modulo two with arbitrary codimension*, Bull. Amer. Math. Soc. **76** (1970), 767–771. MR MR0260981 (41 #5601)

[FF60] Herbert Federer and Wendell H. Fleming, *Normal and integral currents*, Ann. of Math. (2) **72** (1960), 458–520. MR MR0123260 (23 #A588)

[Fle62] Wendell H. Fleming, *On the oriented Plateau problem*, Rend. Circ. Mat. Palermo (2) **11** (1962), 69–90. MR MR0157263 (28 #499)

[Giu84] Enrico Giusti, *Minimal surfaces and functions of bounded variation*, Monographs in Mathematics, vol. 80, Birkhäuser Verlag, Basel, 1984. MR MR775682 (87a:58041)

[GT01] David Gilbarg and Neil S. Trudinger, *Elliptic partial differential equations of second order*, Classics in Mathematics, Springer-Verlag, Berlin, 2001, REPRINT of the 1998 edition. MR MR1814364 (2001k:35004)

[Har79] R. Hardt, *An introduction to geometric measure theory*, Lecture notes given at Melbourne University, 1979.

[Mir67] M. Miranda, *Sulle singolarità delle frontiere minimali*, Rend. Sem. Mat. Univ. Padova **38** (1967), 180–188. MR MR0221351 (36 #4403)

[Pro60] M. H. Protter, *Unique continuation for elliptic equations*, Trans. Amer. Math. Soc. **95** (1960), 81–91. MR MR0113030 (22 #3871)

[Rei60] E. R. Reifenberg, *Solution of the Plateau Problem for m-dimensional surfaces of varying topological type*, Acta Math. **104** (1960), 1–92. MR MR0114145 (22 #4972)

[Rei64] ———, *An epiperimetric inequality related to the analyticity of minimal surfaces*, Ann. of Math. (2) **80** (1964), 1–14. MR MR0171197 (30 #1428)

[Ros84] M. Ross, *Federer's structure theorem*, Research Report, Centre for Mathematical Analysis, Australian National University, 1984.

[Roy88] H. L. Royden, *Real analysis*, third ed., Macmillan Publishing Company, New York, 1988. MR MR1013117 (90g:00004)

[Sim68] James Simons, *Minimal varieties in riemannian manifolds*, Ann. of Math. (2) **88** (1968), 62–105. MR MR0233295 (38 #1617)

[Sim83] Leon Simon, *Asymptotics for a class of nonlinear evolution equations, with applications to geometric problems*, Ann. of Math. (2) **118** (1983), no. 3, 525–571. MR MR727703 (85b:58121)

[Sol82] B. Solomon, Thesis, Princeton University, 1982.

[SSY75] R. Schoen, L. Simon, and S. T. Yau, *Curvature estimates for minimal hypersurfaces*, Acta Math. **134** (1975), no. 3-4, 275–288. MR MR0423263 (54 #11243)

[Whi82] B. White, Thesis, Princeton University, 1982.